REPORT DOCUMENTATION PAGE

AFRL-SR-AR-TR-03-

0231

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE 26 June 2003	3. REPORT TYPE AND DATES COVERED Final Technical Report
4. TITLE AND SUBTITLE Working memory capacity and focused and sustained attention		5. FUNDING NUMBERS F49620-00-1-0131
6. AUTHOR(S) Randall W. Engle and Michael J. Kane		
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Georgia Tech Research Corporation Georgia Institute of Technology Atlanta, Georgia 30302		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Office of Scientific Research Bolling AFB DC, 20332-6448		10. 20030709 017

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12 a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution unlimited.	12 b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words) Results are reported from sixteen different sets of studies on a model of working memory capacity. We see WM as a system consisting of those long-term memory traces active above threshold, the procedures and skills necessary to achieve and maintain that activation and, what we call executive attention – the ability to control and sustain focus of attention. Tasks of working memory capacity (WMC) reflect influences from both domain-specific and domain-free processes but we have concluded that the portion that reflects domain-free executive attention is responsible for the value of such tasks for predicting performance on so many different cognitive measures and is responsible for the relationship between measures of WMC and general fluid intelligence. Our findings suggest that executive attention is important to a wide range of tasks from the realms of social, cognitive, and emotional behavior. Individual differences in executive attention reflect differential functioning of brain circuits in the prefrontal cortex and the anterior cingulate.

14. SUBJECT TERMS			15. NUMBER OF PAGES 43
			16. PRICE CODE
17. SECURITY CLASSIFICATION OR REPORT **UNCLASSIFIED**	18. SECURITY CLASSIFICATION ON THIS PAGE **UNCLASSIFIED**	19. SECURITY CLASSIFICATION OF ABSTRACT **UNCLASSIFIED**	20. LIMITATION OF ABSTRACT UL

NSN 7540-01-280-5500 **Standard Form 298**
(Rev.2-89)
Prescribed by ANSI Std. 239-18
298-102

MASTER COPY: PLEASE KEEP THIS "MEMORANDUM OF TRANSMITTAL" BLANK FOR REPRODUCTION PURPOSES. WHEN REPORTS ARE GENERATED UNDER THE ARO SPONSORSHIP, FORWARD A COMPLETED COPY OF THIS FORM WITH EACH REPORT SHIPMENT TO THE ARO. THIS WILL ASSURE PROPER IDENTIFICATION. <u>NOT TO BE USED FOR INTERIM PROGRESS REPORTS</u>; SEE PAGE 2 FOR INTERIM PROGRESS REPORT INSTRUCTIONS.

MEMORANDUM OF TRANSMITTAL

Dr. Willard Larkin
Program Manager for Life Sciences,
AFOSR/NL
801 North Randolph Street
Arlington, VA 22203-1977

☐ Reprint (Orig + 2 copies)

☐ Technical Report (Orig + 2 copies)

☐ Manuscript (1 copy)

☐ Final Progress Report (Orig + 2 copies)

☐ Related Materials, Abstracts, Theses (1 copy)

CONTRACT/GRANT NUMBER: F49620-00-1-0131

REPORT TITLE: Final Technical Report for Grant - Working memory capacity and focused and sustained attention

is forwarded for your information.

SUBMITTED FOR PUBLICATION TO (applicable only if report is manuscript):

Sincerely,

Randall W. Engle

Final Report

Grant*: F49620-00-1-0131*
Grantee: ***Randall W. Engle***
Title*: **Working memory capacity and focused and sustained attention.***

Unrestricted distribution, unclassified

26 June 26, 2003

Prepared for:

(1) *AIR FORCE OFFICE OF SCIENTIFIC RESEARCH*
BOLLING AFB DC, 20332-6448

Results are reported from sixteen different sets of studies on a model of working memory capacity. We see WM as a system consisting of those long-term memory traces active above threshold, the procedures and skills necessary to achieve and maintain that activation and, what we call executive attention – the ability to control and sustain focus of attention. Tasks of working memory capacity (WMC) reflect influences from both domain-specific and domain-free processes but we have concluded that the portion that reflects domain-free executive attention is responsible for the value of such tasks for predicting performance on so many different cognitive measures and is responsible for the relationship between measures of WMC and general fluid intelligence. Our findings suggest that executive attention is important to a wide range of tasks from the realms of social, cognitive, and emotional behavior. The model assumes that individual differences in executive attention reflect differential functioning of brain circuits in the prefrontal cortex and the anterior cingulate.

The work funded by this grant was about the nature of working memory capacity (WMC), and in this report I will address the nature of WMC limitations, their effects on higher order cognitive tasks, their relationship to attention control and general fluid intelligence, and their neurological substrates. Much of our work has explored these issues in the context of individual differences in WMC and the cause of those individual differences. However, our ultimate goal is to understand WMC in its most general sense. We have used individual differences much in the way suggested by classic papers by Underwood (1975), who urged that individual differences be used as a crucible in which to test theory (see also Kosslyn et al. 2002), and Cronbach (1957), who argued that the two schools of psychology based on experimental and psychometric methods could be synergistic of one another.

We report the status of a nearly two-decade pursuit of the nature and cause of the relationship between "span" measures of WMC and complex cognition. One of the most robust, and we believe, interesting and important findings in research on working memory is that WMC span measures strongly predict a very broad range of higher-order cognitive capabilities, including language comprehension, reasoning, and even general intelligence. In due course, we describe our current thinking about the nature of these relationships and the ramifications for theories of working memory, executive attention, intelligence, and the brain mechanisms underlying those constructs.

Let us first try to place WMC in a context of general theories of immediate memory. In the 1970s and 1980s, after twenty years of work on short-term memory (STM) from the information-processing perspective, many theorists questioned the value of that work, the methods used, and the importance of the findings. Crowder (1982), in a paper pointedly entitled "The demise of short-term memory," argued against the idea that we needed two sets of principles to explain the results of tasks measuring immediate memory and tasks clearly reflecting long-term memory (LTM). He concluded, much as his mentor Arthur Melton did in 1963, that there was insufficient evidence to support the notion of multiple memories. Evidence for a long-term recency effect similar to that found with immediate recall seemed to nullify the relationship between the recency portion of the serial position curve and STM (e.g., Baddeley & Hitch, 1977; Roediger & Crowder, 1976). Studies from the levels-of-processing perspective (e.g., Craik & Watkins, 1973; Hyde & Jenkins, 1973) demonstrated that length of time in storage had little or no impact on delayed recall, contrary to quite specific predictions of the Atkinson and Shiffrin (1968) model. These studies suggested that memory was the residual of perceptual processing of an event and

that orienting tasks that drove different perceptions of the event would lead to different types of codes and, in turn, differential recall. Crowder (1982) also called attention to the fact that individual differences studies had shown an inconsistent relationship between simple STM measures and such complex tasks as reading (Perfetti & Lesgold, 1977). If STM exists and is as important to higher-order cognition as early models suggested – that is, if STM is the bottleneck of the processing system – then one would expect measures of STM to correlate with performance in complex tasks such as reading comprehension.

Baddeley and Hitch (1974) questioned the simple notion of STM on these very grounds, but rather than abandon the notion of an immediate memory that is separate from LTM, they proposed a "working memory" model to supplant STM. Unlike the modal model of STM, working memory theory stressed the *functional* importance of an immediate-memory system that could briefly store a limited amount of information in the service of ongoing mental activity. It is quite unlikely that immediate memory evolved for the purpose of allowing an organism to store or rehearse information (such as a phone number) while doing nothing else. Instead, an adaptive immediate-memory system would allow the organism to keep task-relevant information active and accessible during the execution of complex cognitive and behavioral tasks. The "work" of immediate memory is to serve an organism's goals for action. Baddeley and Hitch were therefore more concerned about the interplay of storage and processing of information than about short-term storage alone. Empirically, they demonstrated that requiring concurrent memory for one or two items had virtually no impact on reasoning, sentence comprehension, and learning. Even when the concurrent memory load approached span length, performance was not devastated as should have been the case if STM was crucial to performance in these tasks. This finding led Baddeley and Hitch to propose separate components of the working memory system that traded off resources in order to handle competing storage and processing functions.

As developed by Baddelely (1986, 1996, 2000), the working memory model now arguably emphasizes structure over function. It consists of both speech-based and visual/spatial-based temporary storage systems (the phonological loop and visuo-spatial sketchpad), with associated rehearsal buffers, as well as an "episodic buffer" thought to maintain episodic information using integrated, multi-modal codes. Finally, a central executive component, analogous to Norman and Shallice's (1986) supervisory attention system, regulates the flow of thought and is responsible for implementing task goals. Much of the experimental and neuroscience research on working memory has been directed at the nature of the phonological loop and visual-spatial sketchpad (Baddeley,

1986; Jonides & Smith, 1997), and although these "slave systems" are easily demonstrated by a variety of lab-based experimental paradigms, their importance to real-world cognition appears to be rather limited in scope (but see Baddeley, Gathercole & Papagno, 1998).

We take a functional approach to the study of immediate memory, which is more akin to the original Baddeley and Hitch (1974) work than to Baddeley's more recent proposals (1986; 1996; 2000; Baddeley & Logie, 1999). Specifically, we emphasize the interaction of attentional and memorial processes in the working memory system, and we argue that this interaction between attention and memory is an elementary determinant of broad cognitive ability. Moreover, we endorse Cowan's (1995, 1999) proposal that the coding, rehearsal and maintenance processes of immediate memory work upon activated LTM traces, rather than retaining separate representations in domain-specific storage structures. As illustrated in our measurement model depicted in Figure 1, STM is represented as activated LTM, and this activation may be maintained or made accessible via a number of strategies or skills (e.g., chunking; phonological rehearsal) that may differ across various stimulus and/or response domains. Attentional, or "executive" processes may also contribute to maintaining access to memory traces if routine rehearsal strategies, such as inner speech, are unavailable, unpracticed, or otherwise unhelpful for the task at hand, or if potent distractors are present in the environment. Our idea is that immediate memory, and executive attention in particular, is especially important for maintaining access to stimulus, context, and goal information in the face of interference or other sources of conflict.

By our view, then, working memory is a system of: (a) short-term "stores," consisting of LTM traces in a variety of representational formats active above a threshold, (b) rehearsal processes and strategies for achieving and maintaining that activation, and (c) executive attention. However, when we refer to individual differences in *working memory capacity* (WMC), we really mean the capability of just one element of the system: executive-attention. Thus, we assume that individual differences in WMC are not really about memory storage per se, but about executive control in maintaining goal-relevant information in a highly active, accessible state under conditions of interference or competition. In other words, we believe that WMC is critical for dealing with the effects of interference and in avoiding the effects of distraction that would capture attention away from maintenance of stimulus representations, novel productions or less habitual response tendencies. We also believe that WMC is a domain general construct, important to complex cognitive function across all stimulus and processing domains.

To better illustrate our view, let us place WMC in a context of general cognition. We believe that much of what we need to know to function, even in the modern world, can be derived from retrieval from LTM – retrieval that is largely automatic and cue-driven in nature. Under those circumstances, WMC is not very important. Even in some putatively complex tasks such as reading, WMC is not required in all circumstances (Caplan & Waters, 1999; Engle & Conway, 1998). However, as we see in the following example, proactive interference can lead to problems from automatic retrieval. When the present context leads to the automatic retrieval of information, which in turn, leads to an incorrect or inappropriate response in a task currently being performed, a conflict occurs between the automatically retrieved response tendency and the response tendency necessary for the current task. That conflict must often be resolved rather quickly, and so we need to have some way to keep new, novel, and important task-relevant information easily accessible.

Take a simple example obvious to every American walking the streets of London for the first time. While driving in a country such as England can lead to potentially dangerous effects of proactive interference, there are numerous cues such as the location of the steering wheel, the cars on your side of the road, etc, prompting the maintenance of the proper task goals. However, in walking the streets of England, the cues are much like those present when walking the streets of any large American city and the temptation – shall we say prepotent behavior – is to look to the left when crossing the street. This can be disastrous. So much so, that London places a warning, written on the sidewalk itself, on many busy cross walks used by tourists. This is a situation in which the highly-learned production, "*if crossing street then look left*," must be countered by a new production system leading to looking to the right when crossing streets. This task seems particularly problematic when operating under a load such as reading a map or maintaining a conversation. For individuals that travel back and forth between England and America, they must keep the relevant production in active memory to avoid disaster.

I. The Measurement of Working Memory Capacity

WMC, the construct, is tied to a sizable number of complex span tasks that we detail below. We describe these in some detail because measures of WMC and STM, like all other measures used by psychologists, reflect multiple constructs or influences. Simple span measures of STM (such as word, letter, and digit span) require subjects to recall short sequences of stimuli immediately after their presentation. We believe that these tasks tell us primarily about domain-specific rehearsal processes, such as inner speech, and domain-specific knowledge, for example pertaining to word meanings or the recognition of salient digit patterns. And, at least among healthy

adults, these simple STM tasks tell us relatively little about executive attention (although we assume that attentional processes play some role even here). In contrast, performance of complex WMC span tasks (such as the operation span, reading span, and counting span), while also relying on speech-based or visual-spatial-based coding, also reflect an individual's capability for executive attention above and beyond domain-specific STM. This is because these tasks require subjects to maintain stimulus lists, in the face of proactive interference from prior lists, while also performing a demanding secondary task. Here, then, stimulus information must remain accessible across attention shifts to and from the processing-task stimuli, thus taxing executive control.

Complex span tasks of WMC were first developed by Daneman and Carpenter (1980), in the context of prior research that failed to find a relation between measures of immediate memory and measures of complex cognition. Daneman and Carpenter reported results from a task that measured memory for short lists of recently presented items and that also showed substantial correlations with a variety of reading comprehension measures. Their reading span task required subjects to read sets of sentences and to recall the last word of each sentence. They defined reading span as the largest set of sentence-final words recalled perfectly. The assumption behind the task was that reading requires a variety of procedures and processes and that those procedures will be more efficient and automated in good readers. Hence, good readers will perform them more efficiently than will poor readers. This, in turn, leaves additional resources available for good readers to store the intermediate products of the comprehension process and for other processes. Thus, in the reading span task, simply reading the sentences aloud and comprehending them would result in differential resources available for storage across subjects. Good readers would have more resources available for storage related processes such as encoding and rehearsal and consequently would recall more sentence-final words. To reiterate, the assumption is that better recall of the words results from better reading-specific skills used to read and comprehend the sentence portion of the task. A simple word span task involving a quite similar demand to the storage portion of the reading span and with similar words should not show a correlation with comprehension measures because the task did not invoke reading-specific processing.

Daneman and Carpenter had subjects perform the reading span task, a simple word span task, and a reading comprehension task consisting of silent reading of 12 passages, averaging 140 words each, with each passage followed by questions about facts or pronominal referents from the passage. In addition, subjects self reported their Verbal Scholastic Aptitude Test score (VSAT). The word span task showed modest but non-

significant correlations with reading comprehension (average .35). However, the reading span correlated .59 with VSAT, .72 with answers to fact questions from the passages, and .90 with answers to questions about the noun in the passage to which a pronoun referred. Further, the relationship between correct pronominal reference and reading span increased as a direct function of the distance between the pronoun and the noun to which it referred. This supported Daneman and Carpenter's contention that people who scored high on the reading span task kept more information active in memory and/or for a longer period of time than did those who scored low on the task.

Daneman and Carpenter (1980; 1983) argued that the substantial correlation between recall on the reading span and measures of comprehension occurs because of individual differences in performing reading-specific procedures during reading. That is, differences on the reading span are caused by differences in residual capacity, in turn, caused by differences in skill at performing reading-specific procedures. If the correlation between the reading span score and reading comprehension occurs because of reading-specific skills and knowledge common to both tasks as Daneman and Carpenter argued, then a complex span task that requires a very different set of skills than reading should not correlate with measures of reading comprehension. By their logic, people have a large reading span score because they are good readers.

Turner and Engle (1989) suggested an alternative view, namely, that people are good readers because they have large working memory capacities independent of the task they are currently performing. They tested a large sample of subjects on four different complex span tasks and two simple span tasks. Two tasks were modeled after the reading span. The sentence word task was identical to reading span except half the sentences were nonsense and subjects had to decide whether each sentence made sense and they recalled the sentence-final words. In the sentence digit task, subjects read and made decisions about sentences but instead of remembering the last word, they recalled a digit that occurred after each sentence. In the operation spans, subjects saw and read aloud an operation string such as 'Is (9/3) – 2 = 1?' They were to say yes or no as to whether the equation was correct. In the operation-digit span task, they were to recall the digit to the right of the equal sign for each operation in the set. In the operation-word span task, they were to recall a word that appeared to the right of the question mark. Thus, half the tasks involved reading sentences and half involved solving arithmetic strings. Half involved recalling digits and half involved recalling words. In addition, subjects received a simple word span and simple digit span task. As measures of comprehension, Turner and Engle

tested subjects on the Nelson-Denny Reading Comprehension Test and obtained their Scholastic Aptitude Scores from university records. The Daneman and Carpenter view predicts that only those tasks requiring reading would correlate with the comprehension measures. If, on the other hand, working memory capacity is an abiding characteristic of the person, relatively independent of the particular task, then the complex span tasks might correlate with comprehension regardless of whether they involved reading sentences or performing arithmetic.

The results showed that all four of the complex span tasks predicted reading comprehension and the correlations involving the operation spans were actually a bit higher than those tasks requiring reading sentences. Neither of the simple span tasks correlated with comprehension. The complex span tasks clearly reflect some construct important to comprehension that is not reflected in the simple span tasks. However, whether the tasks involve reading sentences or solving arithmetic does not appear to be important. Another analysis performed by Turner and Engle is notable. One possible explanation for the results is that they reflect a spurious correlation between verbal and quantitative skills. That is, people who are good readers may also be good at solving arithmetic and this could provide the results obtained by Turner and Engle but for reasons commensurate with the Daneman and Carpenter argument. However, when the Quantitative SAT was partialled out of the correlation between the span tasks and comprehension, the operation word span remained a significant predictor of comprehension, and, indeed, the operation word span contributed significant variation in comprehension even after the effects due to the sentence word span were eliminated. These findings led Turner and Engle to conclude that 'Working memory may be a unitary individual characteristic, independent of the nature of the task in which the individual makes use of it.' (pg. 150).

A. Validity of the Relationship

If the measures of working memory capacity are valid measures of a construct with wide ranging importance, then the measures should correlate with a wide range of other cognitive measures and that is indeed the case. We provide below a partial and evolving list of tasks that correlate with measures of WMC. This list is particularly impressive given the notable lack of such relationships with simple span measures of temporary memory (Dempster, 1981).

We view WMC as an abiding trait of the individual, resulting from differences in the functioning of normal brain circuits and neurotransmitters. We see WMC as a cause of inter-individual differences in performance of a huge array of cognitive tasks where the control of attention is important. However, intra-individual reductions in

capability for attention control can also be a result of many different conditions from drunkenness to fatigue; from damage to the frontal lobe to psychopathology. It is becoming clear that conditions such as depression (Arnett et al., 1999), post-traumatic stress disorder (Clark et al., 2003), and schizophrenia (Barch et al., 2003), lead to reductions in WMC even when measures of STM show no decrement. Thus, studies of the results of individual differences in WMC should enlighten us about cognition in these other conditions as well.

Scores on WMC tasks have been shown to predict a wide range of higher-order cognitive functions, including: reading and listening comprehension (Daneman & Carpenter, 1983), language comprehension (King & Just, 1991), following directions (Engle, Carullo, & Collins, 1991), vocabulary learning (Daneman & Green, 1986), note taking (Kiewra & Benton, 1988), writing (Benton, Kraft, Glover, & Plake, 1984), reasoning (Barrouillet, 1996; Kyllonen & Christal, 1990), bridge playing (Clarkson-Smith & Hartley, 1990) and computer-language learning (Kyllonen & Stephens, 1990; Shute, 1991). Recent studies have begun to demonstrate the importance of WMC in the domains of social/emotional psychology and in psychopathology, either through individual differences studies or studies using a working memory load during the performance of a task (Feldman-Barrett et al., in press). For example, high WMC subjects are better at suppressing thoughts about a designated event (Brewin and Beaton, 2001). Likewise, low WMC individuals are less good at suppressing counterfactual thoughts, that is, those thoughts irrelevant to, or counter to, reality. We have also made the argument (Engle, Kane, & Tuholski, 1999) that attentional-load studies are a valuable technique to study intra-individual differences in WMC since a secondary attentional load would reduce WMC. For example, Goldinger et al (2003) found that low WMC subjects showed more counterfactual thinking than did high WMC subjects, but only under conditions of a secondary load. In the absence of a load, there was no difference between high and low WMC subjects since both groups could presumably control their counterfactual thoughts.

Richeson and her colleagues (Richeson & Shelton, 2003) have argued that WMC comes into play in the regulation of automatically activated prejudicial attitudes. White subjects were given a test of implicit attitudes, and then interacted with a white or black 'partner', before performing the Stroop task. The argument was that individuals whose implicit attitude showed them to be more prejudicial against blacks would have to use more of their WMC to block their attitudes while interacting with a black partner than with a white and should do worse on the subsequent Stroop task. That is what Richeson and Shelton (2003) found. Whites who scored high on

prejudice on the attitude test, did worse on the Stroop after interacting with a black partner than when they interacted with a white partner.

WMC has also been used in explanations of various psychopathologies. For example, Finn (2002) proposed a cognitive-motivational theory of vulnerability to alcoholism and one of the key factors is WMC. He argues that greater WMC allows an individual to better monitor, manipulate, and control behavioral tendencies resulting from personality characteristics and that this directly affects the ability to resist a prepotent behavior such as taking a drink in spite of being aware that such behavior is ultimately maladaptive.

Measures of WMC also appear to have some utility as diagnostic measures in neuropsychology. Rosen and her colleagues (Rosen et al., 2002) tested two groups of middle-aged individuals, one of whom consisted of individuals who were carriers of the e4 allele associated with early onset Alzheimer's disease, and the other consisting of non-carriers of the allele. Even though the carriers showed no symptoms of Alzheimer's disease and very few other cognitive measures distinguish between the two groups, the e4 carriers performed significantly worse on the operation span task than did controls. This suggests that operation span, and likely other WMC measures as well, reflect a construct that is unusually sensitive to early changes associated with Alzheimer's.

The wide range of tasks and conditions associated with performance on WMC measures suggests that tasks such as operation and reading span are valid measures of a construct that is an important component of complex cognition reflective of neurological function, thus showing good construct validity. However, as we will see below, WMC is not important to *all* cognitive tasks; the measures also reflect good and lawful discriminant validity. As we will argue below when we discuss our studies using structural equation modeling, this suggests WMC to be a single construct reflecting a domain-free ability for maintaining information in a highly active, easily retrievable state, particularly under conditions of endogenous or exogenous interference.

B. Reliability of the Measures of WMC

Another important characteristic of tasks used to study individual differences is reliability. Experimental psychologists often think of reliability as the likelihood that a phenomenon will replicate from one study to the next as opposed to being due to random fluctuation. Psychometricians think of reliability in terms of whether individuals will show a similar pattern of performance on a given measure from one time to the next. Since our studies often use extreme-groups designs, we are concerned about whether a difference or non-difference found between high and low WMC subjects will replicate across studies. However, we are also concerned about

whether performance on a given WMC task shows strong test-retest correlations with identical or similar forms of the task, as well as whether WMC span tasks are multiply determined.

Reliability is affected by several variables. One that is particularly problematic is the range of the measure. As we will see below, WMC at the construct level is strongly related to general fluid intelligence. Thus, studies using a sample from a highly selected university population will likely have a very restricted range of true-score WMC and the reliability of the measures will be reduced substantially under those conditions. Likewise, extreme-groups designs that use a median split to define high and low WMC subjects are likely to be insensitive to true-score differences in the groups and would need quite large samples to replicate findings from extreme-groups studies using upper and lower quartiles to define the groups.

Reliability of WMC measures has been measured in several ways. One is the internal consistency of the measures, normally done with split-half correlations known as coefficient alphas. Alphas for WMC measures are rarely as low as .7 and are often in the .8 – .9 range. In other words, half the test will correlate with the other half the test in that range (Engle et al 1999; Turner & Engle, 1989). The other way reliability has been assessed is to calculate the correlation between scores on the task from two or more administrations. Klein and Fiss (1999) tested a sample of subjects on the operation span task, and then tested them again after three weeks on an equivalent form of the task, then tested them again 6 – 7 weeks later. They found a corrected reliability estimate of .88 across the three administrations. They also found the rankings of individuals from time one to time two to time three to be quite similar. Thus, the operation span task appears to be highly reliable and quite stable across time. Such extensive analyses has not been performed for the reliability of other WMC measures but we would expect them also to be quite high if the sample of subjects is not highly restricted on general ability measures.

II. Alternative Explanations of the WMC × Higher Order Cognition Correlation

Measures of WMC are reliable and valid, but what are the psychological mechanisms responsible for the fact that they correlate with such a wide array of higher-level cognitive tasks? First, we need to make a methodological point here that is probably obvious but needs to be stated. We need to constantly remind ourselves about the difficulty of attributing cause-effect relationships in psychology. Further, all readers will certainly understand the difficulty of attribution about cause and effect when describing a correlation between two variables. Daneman and Carpenter reported, at base, a correlation between a span measure and one or more measures of comprehension. Turner and Engle showed that the explanation for the correlation given by

Daneman and Carpenter was inadequate. However, the question as to what causes a correlation is a tricky one to answer and just about everything else we describe in this paper was done in pursuit of an answer to that question. The difficulty, of course, is that some third variable, bearing little direct relationship to either of the two measures, might drive the putative relationship between the two observed variables. Our strategy for understanding the nature of this correlation takes a two-pronged approach very much following Cronbach's (1957) advice about the two schools of psychology, one experimental and the other psychometric. One approach, referred to as *microanalytic* (Hambrick, Kane & Engle, in press), has been to treat the correlation as a dependent variable and to perform experimental manipulations testing various hypotheses to see whether the correlation between working memory capacity (WMC) measures and higher-order cognition is affected. The presumption is that if we can make the correlation appear and disappear with a given manipulation, some aspect of the manipulation controls the correlation. A typical experiment uses an extreme-groups design with subjects from the upper and lower quartiles on one or more WMC measures, with the test being whether high and low WMC subjects perform differently on some cognitive task. For example, a study showing that high and low WMC subjects differ on a version of a task under conditions of proactive interference but do not differ on a version of the task absent the interference is suggestive that interference might play a role in the nature of the correlation.

The other approach, referred to as *macroanalytic* (Hambrick et al., in press), is to test a large number of subjects on a large number of tasks representing various constructs and perform structural equation modeling to determine the relationship among various constructs. The first approach is cheaper and quicker to determine whether individual differences in WMC are important to a task and the variables that interact with WMC in that task. It allows subtle manipulations in tasks that would be prohibitive using the second approach. However, one cost is that it over-estimates the degree of relationship between the two variables. The second approach is more expensive in time and labor but gives a much cleaner and clearer picture of WMC at the construct level and the degree of relationship of other constructs with WMC.

The following alternative explanations have been suggested, but as will be seen, have not been supported by the evidence.

A. Word Knowledge

We have used both approaches, sometimes in the same study, to investigate potential explanations for the correlation. For example, Engle, Nations, and Cantor (1990) tested the idea that the correlation between the

span and comprehension measures occurs because of individual differences in word knowledge. Complex span measures requiring recall of words typically are more predictive of comprehension than those requiring recall of digits (Daneman & Merickle, 1996; Turner & Engle, 1989), thus, the correlation could be a spurious one involving word knowledge. People who know more words and more about words will be more familiar with the words in span tasks and in text passages and will score higher on both types of tasks. If that explanation were correct, then the span-comprehension correlation should be high when the span task requires retention of low frequency words, because word knowledge would be more variable across subjects, but low when very high frequency words are used since word knowledge should not differ that much across subjects. Engle et al. (1990) tested 90 subjects representing a rectilinear distribution of the Verbal SAT range on simple and operation span tasks using low and high frequency words. The question was whether comprehension, as represented by the VSAT, would correlate with span measures with both high and low frequency words. The answer was yes, for the complex span measures, both low and high frequency words equally predicted VSAT. Thus, the idea that variation in word knowledge is the third variable responsible for the correlation between complex span and comprehension is not supported.

Engle, Cantor and Carullo (1992) reported a test of other alternative explanations of the WMC correlation with higher-order cognitive tasks. We first describe the methodology, then the various explanations, and then describe the results pertinent to each of the possible explanations. In one experiment, subjects performed a self-paced version of the operation span task and, in a second experiment, the reading span task. Both used a moving-window procedure to present each element of the operation or sentence and the to-be-remembered word. Key-press times were used as an estimate of processing efficiency for the processing portion of the task and for the amount of time subjects spent studying the to-be-remembered word following either the operation or the sentence. For example, to show the operation-word string "(6/2)-1=_____. knife", the first key-press would present an open parenthesis and a single digit { (6 }, the second key-press would turn off the first display and present either a multiplication or division sign { / }, the third would present a single digit and a close parenthesis { 2) }, the next press would present an addition or subtraction sign { - }, next a single digit (1 }, next an equal sign and underscore line { =____ }, the subject then typed in the single digit answer, and the word { knife } was shown until a key press started the next string.

Subjects first performed a series of the operations without recalling the word and in the other experiment, with reading span, simply read the sentences. The time between key presses was measured as an index of the processing efficiency for the elements of the processing portion of the task. Subjects then performed the operation span task with sets of two to six items and recall of the words from that set afterward. Again, processing times were recorded for the elements of the display including the time that subjects spent looking at the words to be recalled. Reading comprehension was measured by the Verbal Scholastic Aptitude Test.

B. Task specific hypothesis.

This view, the original explanation advanced by Daneman and Carptenter (1980), is that the correlation between a measure of higher-order cognition and a measure of WMC will only occur if the processing portion of the WMC task requires the same skills and procedures as the higher-order task. If that explanation is correct, we should see a correlation between the time to view the sentence words of the reading span task, words recalled in the reading span task, and VSAT. Note that these relationships should hold for the processing task *without recall* as well as the reading span task *with* recall, since it is based on skill at performing the processing portion of the task. However, the relationship should not hold for the operation span task because the processes required to solve the equations are unlikely to be similar to those used in reading the passages for the VSAT.

C. General processing hypothesis.

This view, representing the thinking of Case (1985), argues that individual differences in WMC occur because some people do all mental operations faster and more efficiently than others do. Thus, reading and arithmetic operations both would be done faster and more efficiently, leading to greater residual resources for storage of the to-be-remembered words. If this hypothesis is correct, then the correlation between number of words recalled in both the reading span and the operation span and VSAT should be significant. However, it also predicts a correlation between the viewing times for the elements of the arithmetic and reading portions of the task and the number of items recalled in the span task. Further, this relationship between element viewing times and recalled items should hold even for viewing the elements in a task *without* recall. In addition, if we partialled out the variance attributable to viewing the elements, from either the tasks with or without recall, from the span/VSAT correlation, that correlation should be eliminated or at least significantly reduced.

D. Strategic allocation hypothesis

This view is an extension of the ideas reported in Carpenter and Just (1989). They suggested that high spans better allocate their resources between the processing and storage portions of the task than do low spans. That is, as load increases, high spans redirect resources away from the processing portion to the increasing storage element. Low spans do not adjust their resource allocation strategy as load increases. If this explanation accounts for the greater recall in complex span tasks by high span subjects, then we should see that high spans spend less and less time viewing the elements of the processing portion of the task as load increases. Further, there should be a negative correlation between processing time and number of span words recalled. Additionally, if we partialled processing times out of the span/VSAT relationship, the correlation should be eliminated or reduced. These predictions should hold for both operation span and reading span.

E. Rehearsal differences hypothesis.

The idea behind this hypothesis is that the correlation between WMC scores and higher order cognition occurs because some high WMC individuals are more likely to rehearse in the span tasks and also to be more strategic in other tasks as well. According to this hypothesis, there should be a positive correlation between time spent viewing the to-be-remembered words in both operation and reading span and the number of words recalled. More importantly, however, partialling out the time spent studying the to-be-remembered words from the span/VSAT relationship should eliminate or reduce the correlation.

The Engle et al (1992) results were quite clear in eliminating all of these hypotheses. First, replicating Turner and Engle (1989), the number of words recalled in both operation span and reading span significantly correlated with VSAT and at the same level. Secondly, processing times on the storage-free versions of the task did not distinguish between high and low WMC individuals. Time spent viewing the elements did not consistently correlate with the span score. Thirdly, when the processing times for the elements of operation and reading spans, both with and without recall, were partialled out of the span/VSAT correlation, the correlation was not diminished. In fact, there was a slight trend for the correlation between operation span and VSAT to go up. Fourthly, there was a significant correlation between viewing time of the to-be-remembered words and the span score, with high spans spending more time viewing the words than did low spans. However, when those times were partialled out of the span/VSAT correlation, the correlation was unchanged.

This suggested to us that individual differences in rehearsal time did affect the number of words recalled in this task, but that this is a *nuisance variable* unrelated to the construct responsible for the relationship between WMC and reading comprehension. This issue merits further discussion since it is apparently misunderstood in the literature. For example, McNamara and Scott (2001) demonstrated that strategy training led to an increase in scores on a WMC span task. From that, they concluded that the correlation between span and higher-order cognition was a result of differences in strategy use with high WMC subjects more likely to use strategies than low spans. We have repeatedly made the point (Engle et al 1999) that the complex span score, like all cognitive measures, is a result of a multitude of constructs and that manipulations may affect some contributors to the score while having no impact on the construct mediating the score and the vast array of higher-order cognitive tasks. As Engle et al (1992) showed, subjects who studied the to-be-remembered word longer on the operation span and reading span had higher span scores. However, study time did not contribute to the relationship between span and VSAT. Many different variables would lead to better or worse performance on WMC tasks such as operation span and reading span. However, the critical question is whether those same variables eliminate or reduce the correlation between the span score and measures of higher-order cognition such as reading comprehension or spatial reasoning. That is the only way to determine whether the variable is important to an explanation of the correlation. Thus, although McNamara and Scott demonstrated that training a particular strategy may increase span scores overall, they did not demonstrate that strategies are at all related to the processes that link WMC to complex cognition. In fact, one may infer that their strategy training actually *increased* individual differences in complex span, rather than reduced them, as the standard deviations in span were slightly larger after training than before, especially for subjects who were less strategic originally. These findings thus leave open the possibility that strategy training benefits some individuals more than others, with the degree of this benefit tied to WMC, thus reversing causal inference made by McNamara and Scott.

A more direct test of the rehearsal or strategy differences hypothesis was made by Turley-Ames and Whitfield (2003). Their study measured a large number of subjects (n=360) on the operation span task who were then assigned to either a no-training control group, rote rehearsal group, imagery strategy group, or semantic association group similar to McNamara and Scott's chaining condition. All subjects were retested on the operation span and then the Nelson-Denney Reading Comprehension test. If the correlation between operation span and comprehension results from differences in rehearsal, then training should eliminate or reduce the

correlation between the second operation span and Nelson-Denny. However, if Engle et al (1992) were correct in arguing that rehearsal differences do occur and are important to span score, *but*, they are a nuisance variable with no causal influence, then procedures designed to encourage subjects to behave more similarly with respect to rehearsal strategy should not reduce the span/comprehension correlation. In fact, such procedures should *increase* the correlation by reducing error variance resulting from the nuisance variable. Turley-Ames and Whitfield (2003) found that strategy training was effective in increasing the operation span scores, compared to the control group. However, the correlation between the operation span and Nelson-Denny was higher after strategy training (rote rehearsal r=.56, imagery r=.32, and semantic association r=.47) than in the control condition (r=.30). Thus, differential rehearsal and strategy-use do not account for the correlation and, in fact, appear to serve as a suppressor variable for the true relationship between the span score and higher-order cognition.

Complicating the picture of the relationship between rehearsal and WMC is that greater WMC apparently leads to greater benefit from rehearsal and encoding strategy use, as we foreshadowed previously. Pressley, Cariglia-Bull, Deane, and Schneider (1987) tested children who heard concrete sentences they were to learn. Half the children received instruction in how to construct images representing the sentences. In addition to the sentence-learning task, children also completed a battery of short-term memory tasks including simple word span. Pressley et al found that, while STM capacity was not related to performance in the control condition, it did predict sentence learning quite highly in the strategy learning group, even with age held constant. These results suggest that children with greater WMC may be better able to learn and/or use strategies for learning and retrieval of information. (Note, again, that the causal path implied here is from greater WMC to greater strategy effectiveness and not from greater strategy use to greater WMC.)

F. Speed Hypothesis

Another explanation for the covariation of WMC tasks and other cognitive tasks is that both reflect individual differences in speed of processing. This is a variant of a hypothesis popular in explaining the effects of aging on cognition called "age-related slowing" (Kail & Salthouse, 1994; Salthouse, 1996); it is also similar to views advocated by some theorists of intelligence (Jensen, 1982, 1998). The idea behind age-related slowing is that elemental cognitive processes become slower as we age and this slowing has a ubiquitous, deleterious effect on higher-order cognitive functioning. Thus, the argument goes, low WMC individuals are simply slower to

process all information, and this leads to lower scores on complex WMC measures (perhaps because slowing allows for greater trace decay) and lower scores on other cognitive measures as well.

Many studies in the literature do, in fact, report reasonably strong correlations between processing-speed and WMC constructs (Ackerman, Beier & Boyle, 2002; Kyllonen, 1993; Kyllonen & Christal, 1990; Oberauer, Süβ, Schulze, Wilhelm & Wittmann, 2000; Park, Lautenschlager, Hedden, Davidson, Smith & Smith, 2002; Salthouse & Meinz, 1995). The question is what to make of these correlations. We believe many of them to be artifactual. For example, some studies tested an age range from young adults to elderly adults (Park et al., 2002; Salthouse & Meinz, 1995), and speed need not have the same relation to WMC within an age group, such as young adults, as it does across age groups (see Salthouse, 1995). More worrisome, however, is the fact that in some studies the WMC tasks were presented under time pressure at either study or test (Ackerman et al., 2000; Oberauer et al., 2000). Obviously, presenting subjects with a speeded WMC test will artificially inflate correlations between WMC and "processing speed" measures. In some studies, moreover, the "speed" tasks were quite complex, for example requiring task-set switching, mathematical operations, or the association of arbitrary codes to individual items (Ackerman et al., 2000; Kyllonen, 1993; Kyllonen & Christal, 1991; Oberauer et al., 2000). Although such complexity is desirable because it increases variability and allows correlations to occur, a task analysis of these complex speed tasks strongly suggests that they tax executive attention, immediate memory, and/or LTM retrieval processes (see Conway, Cowan, Bunting, Therriault & Minkoff, 2002; Conway, Kane & Engle, 1999). Given our view that WMC measures fundamentally tap an attention-control capability, causal inferences regarding correlations between WMC and complex speed measures are highly ambiguous – it is just as likely that WMC differences lead to speed differences as is the reverse.

On the logic that WMC and speed measures should be as unconfounded as possible, Conway et al. (2002) tested their subjects in complex span tasks that were untimed, as well as in relatively simple processing-speed tasks. The speed tasks involved making same-different judgments about individual pairs of verbal and non-verbal stimuli, or copying visual lists of digits or letters. Despite their simplicity, these speed tasks yielded substantial variability in the sample. However, Conway et al. found very weak correlations between WMC and speed measures, and furthermore, only the WMC tasks correlated significantly with fluid intelligence. Speed measures did not. A structural equation model clearly demonstrated that processing speed did not account for the relationship between WMC and general cognitive ability.

In our own laboratories, we recently began testing high and low WMC span subjects in attention-control tasks (for a full discussion see below). Important for present purposes is that we typically fail to find RT differences between span groups in the baseline conditions that assess relatively automatic processes (Kane, Bleckley, Conway & Engle, 1999; Kane & Engle, 2003). If low-level processing-speed mechanisms were responsible for WMC differences, then span differences in baseline speed would be expected. Indeed, we have also failed to find span differences in RTs in some fairly complex and difficult tasks such as visual search, even with large arrays of distractors that share perceptual features with the target. As we will discuss below, findings of independence between WMC and "controlled" visual search appear to present boundary conditions on the relationship between WMC and attention control, but here they serve to reinforce the idea that WMC differences cannot be explained merely by variation in "processing speed."

G. Mental Effort/Motivation

Another alternative to the explanation we offer here is that differences in motivation mediate the WMC x higher-order cognition relationship. That is, some individuals are simply more motivated than others to do well on tasks of all types, including complex working-memory tasks and tasks of higher-order cognition. There are four lines of logic against this argument. First, quite lawfully, we find differences between high and low WMC individuals on tasks that require the control of attention but do not see differences in tasks that can be thought of as automatic. As we will describe below, span does not predict performance in the prosaccade task, which depends on a relatively low-level attention capture. We do observe differences, however, on the antisaccade task, which requires that the attentional capture by an exogenous cue be resisted in order to make the correct response of looking to a different region of space (Kane et al., 2001; Unsworth, Schrock, and Engle, 2003). WMC differences are not observed in speed to count objects where the number is within the subitizing range of 1 – 3, but substantial differences are observed when counting a larger number of objects (Tuholski, Engle & Baylis, 2001).

Second, we see WMC differences on memory tasks involving a high level of proactive or retroactive interference but not on the same tasks in the absence of interference. For example, high and low span subjects do not differ on the fan task unless there is overlap among the propositions (Bunting et al, 2002; Cantor & Engle, 1993; and Conway and Engle, 1994). Further, Rosen and Engle (1998), and Kane and Engle (2001) found that low span subjects are much more vulnerable than are high spans to the effects of interference. However, in the

absence of interference conditions, high and low span subjects do not differ, despite the fact that their performance was well off ceiling and floor. We will describe these studies in more detail below, but for now the WMC equivalence in demanding but low-interference memory contexts is difficult to reconcile with motivation explanations for WMC effects.

Third, a motivation explanation must argue that differences between high and low WMC subjects on other tasks should increase as the task becomes more difficult or complex (i.e., as it becomes more effortful). We have observed two strong counterexamples of this prediction, however, in studies not originally directed at the motivation explanation. In one, discussed above in regards to processing speed, we have studied visual search in three different experiments with high and low span subjects (Kane, Poole, Tuholski & Engle, 2003). In all of these studies, subjects searched for a target letter F. Stimulus arrays consisted of few (0 – 3), several (8 – 9), or many (15 – 18) distractors, and these distractors were either dissimilar or similar to the target ("O"s versus "E"s, respectively). As clearly seen in Figure 2, high and low WMC subjects performed identically in both the more "automatic" and the more "controlled" search conditions, despite massive RT increases from small to large stimulus arrays across studies.

We have found similar results in studies of WMC and task-set switching (Kane & Engle, 2003). Three experiments used a numerical Stroop task (Allport, Styles & Hseih, 1994), in which subjects were cued unpredictably to either switch between counting arrays of digits and reporting the digits' identity, or repeat the same task with consecutive arrays. A fourth experiment, with four between-subject conditions, used a letter/digit judgment task (Rogers & Monsell, 1995), in which subjects predictably repeat and switch tasks in an AABB task sequence. We found the typical switch cost, i.e., the RT difference between task-switch and task-repeat trials, in all experiments. However, in *no* experiment did we find any span difference in switch costs despite the fact that overall switch costs were robust. Clearly, a motivation explanation cannot account for the absence of span differences in demanding search and switching tasks. Indeed, in one of our Stroop switching experiments, subjects were allowed to study the task cues for the upcoming trial pair (e.g., "DIGIT → COUNT") for as long as they wanted, and low spans actually studied the cues for significantly *more* time than did high spans, and this span difference was especially pronounced on switch trials. Such extra effort on the most difficult trials is certainly not expected from an unmotivated sample.

Fourth and finally, a series of studies by Heitz et al (2003) used pupil dilation as a measure of mental effort to directly address the contribution of motivation to WMC effects. Pupil dilation has proven to be a sensitive and reliable index of the mental effort allocated to cognitive tasks, with pupil size tending to increase as a task becomes more and more difficult (Kahneman & Beatty, 1966). The motivation explanation argues that performance differences between high and low span subjects results from low spans being poorly motivated relative to high spans. Hence, a manipulation that increases motivation should lead to low spans performing more like high spans. If, on the other hand, high and low WMC subjects are similar in their motivational level, the motivation-enhancing manipulation should lead to similar performance increases for both groups.

Heitz et al. (2003) had subjects who had been selected as high and low span, on the basis of the operation span task, subsequently perform the reading span task under conditions designed to manipulate motivation. In addition to measuring performance on the reading span task, we measured pupil size. In one study, high and low span subjects were provided a financial incentive for performance on the reading span task. They could make up to $20 depending on their recall of letters that followed the to-be-read sentences and on their ability to answer questions about the sentences. The incentive manipulation led to an equivalent increase in reading-span performance for high and low span subjects; that is, both high and low span subjects improved their observed "span" with incentives, but the difference between the two WMC groups remained unchanged. In addition, the incentive manipulation increased baseline pupil size taken before the beginning of each trial. However, again, the increase was the same for high and low span subjects. Pupil size clearly reflected level of mental effort in the task because pupil size closely mirrored memory load in the reading span task. For example, as a 5-item set progressed from item 1 to 5, pupil size increased for both groups. However, the increase in pupil size was, again, identical for high and low span subjects. It is clear that Heitz et al. successfully manipulated motivation. And, it is equally clear that the lack of differential incentive effects between high and low span subjects means that performance differences related to WMC do not result from generic motivation differences.

III. Macroanalytic Studies of Working Memory Capacity: Its Generality and Relation to other Constructs

Our large-scale, latent-variable studies have addressed questions about WMC at the construct level. Specifically, these studies have assessed the relationship between WMC and other constructs such as STM and general fluid intelligence, and they have also tested whether WMC should be thought of as a unitary, domain-general construct or whether separate verbal and visuo-spatial WMC constructs are necessary.

Before discussing this research in more detail, however, let us briefly note the advantages of latent-variable approaches to the study of WMC. Latent-variable procedures require that each hypothetical construct be measured by multiple tasks (such as using operation span, reading span, and counting span to measure WMC) and they statistically remove the task-specific error variance associated with the individual, multiply determined tasks. What remains, then, is only the variance that is shared among all the tasks, which putatively represents the latent construct of interest, free of measurement error. These statistical methods are valuable because no single task is a pure measure of any one single construct. Operation span, for example, measures not only the latent construct of WMC, but also some degree of math skill, word knowledge, and encoding and rehearsal strategies. Therefore, construct measurement that is based on multiple tasks that differ in their surface characteristics will be more valid than that based on single tasks, which can never be process pure. Latent-variable techniques used with correlational data are therefore analogous to the converging-operations approach in experimental research, in which constructs are validated through multiple and diverse experimental conditions that eliminate alternative hypotheses (Garner, Hake & Eriksen, 1956; see Salthouse, 2001).

Recall that we have portrayed working memory as a system consisting of domain-specific memory stores with associated rehearsal procedures and domain-general executive attention. Engle, Tuholski et al. (1999) tested that idea using an approach by which we identified latent variables through structural equation modeling and determined the relationship among those latent variables. We reasoned that all span tasks are mediated by multiple latent variables. For instance, simple STM tasks such as word, letter, and digit span are verbal tasks, and so they reflect variance due to differences in verbal knowledge and experience with the particular item types. In addition, performance on these tasks is affected by individual differences in pattern recognition (in the case of digit strings) and the frequency and type of rehearsal strategies used. To the extent that such strategies are less well practiced or routinized, one would also expect some contribution of attention control to successful performance.

Complex WMC tasks such as reading span, operation span, and counting span also require retention and recall of words, letters and digits, and so they also reflect variance attributable to these variables. However, we also reasoned that WMC tasks principally reflect individual differences in ability to control attention, due to the demand to maintain items in the face of attention shifts to and from the "processing-task" stimuli. If that were true, then the two types of tasks (WMC and STM) should reflect different – but correlated – latent variables. Moreover,

when we extract the variance common to the two constructs, the residual, unique variance from WMC should reflect individual differences in the ability to control attention. We also tested the idea, proposed by Kyllonen & Christal (1990), that WMC is strongly associated with general fluid intelligence (gF). If that were true, then the WMC construct should be strongly associated with gF, but the STM construct should not. Further, the residual variance from WMC that remains after extraction of a 'common' variable from WMC and STM, representing executive attention, should be strongly associated with gF.

We used three measures of WMC: reading span, operation span, and counting span; three measures of STM: forward word span with dissimilar sounding words, forward word span with similar sounding words, and backward word span; and two measures of gF: Ravens Progressive Matrices (Raven, Raven & Court, 1998) and Cattell Culture Fair Test (cite). Figure 3 shows that a model with separate factors for WMC and STM fit the data quite well and better than a single factor representing all six span tasks. Clearly, the two factors are strongly associated (.68) as we expected, but two factors provided the best fit of the data. You also see from Figures 3 and 4 that, while the link between WMC and gF is quite strong, once the association between WMC and STM is accounted for there is no significant association between STM and gF. In other words, any association that STM tasks such as digit and word span have with fluid abilities occurs because of the strong association STM has with WMC.

Figure 4 shows what happens when the variance common to the two memory constructs is extracted to the latent variable labeled as 'common'. The curved lines represent the correlation between the residuals for WMC and STM and gF, that is, the correlation between each construct and gF after extracting the variance that was shared between WMC and STM tasks. The correlation between gF and the residual variance remaining in WMC after Common was extracted was high and significant (.49). However, the similar correlation between gF and the residual for STM was not significant. This supports the notion that some aspect of WMC other than STM is important to fluid intelligence and presumably to other aspects of higher-order cognition as well. We argue that that critical aspect of WMC tasks is the ability to control attention. This follows from the logic that, if the working memory system consists of STM processes plus executive attention, then after Common is extracted, this should leave executive attention as residual. Of course, there was no direct evidence for this inference by Engle, Tuholski et al. (1999) but we will provide ample evidence to support that conclusion below.

In a more recent large-scale study (Kane, Hambrick, Tuholski, Wilhelm, Payne & Engle, 2003), we have also addressed the question of how much shared variance exists between verbal and visuo-spatial WMC – that is, is it necessary to posit separate latent variables for verbal and spatial complex span tasks, or instead should WMC be considered an entirely domain-general construct? The latter, domain-general hypothesis most easily follows from our view that individual differences in WMC correspond to individual differences in general attentional capabilities. Although there is little doubt that verbal and visual/spatial information are coded differently and by apparently different structures in the brain (Jonides & Smith, 1997; Logie, 1995), a separate question is whether what we have referred to as executive attention must also be fractionated for verbal and visual/spatial formats. Our belief is that executive attention is general across representation formats and is common to both verbal and spatial tasks requiring the control of attention. However, Engle et al (1999) used only verbal tasks, which did not allow us to address this issue.

In conflict with our view, several correlational studies have, in fact, suggested that verbal and visuo-spatial WMC may not only be separable, but also virtually independent (Daneman & Tardif, 1987; Friedman & Miyake, 2000; Handley, Capon, Copp & Harper, 2002; Morrell & Park, 1993; Shah & Miyake, 1996). All of these studies presented university students with one complex span task using verbal materials and one complex span task using visuo-spatial materials, and these WMC tasks were used to predict some higher order verbal and visuo-spatial task (or task composite). In short, the verbal and visuo-spatial span tasks were poorly to modestly correlated with one another, and each correlated more strongly with complex cognition in its matching domain than in the mismatching domain: Verbal span predicted verbal ability better than spatial ability, and spatial span predicted spatial ability better than verbal ability. Indeed, the correlations for mismatching span and ability tasks were typically non-significant and often near zero.

Nonetheless, we had good reason to doubt that WMC was primarily or entirely domain-specific. First, the breadth of predictive utility demonstrated by verbal WMC tasks, including their strong correlations with non-verbal tests of fluid intelligence (Conway et al., 2002; Engle, Tuholski et al., 1999) and their relation to rather low-level attention tasks (to be discussed below; Conway, Cowan & Bunting, 2001; Kane et al., 2001; Kane & Engle, 2003; Long & Prat, 2002) indicates that verbal WMC tasks tap something important beyond just verbal ability.

Second, the studies that indicated domain specificity had methodological problems that could have systematically led to an underestimation of WMC's generality. Most obviously, some of the verbal and visuo-

spatial tasks differed markedly in their difficulty, making their discrepant patterns of correlations impossible to interpret (Daneman & Tardif, 1987; Morrell & Park, 1993). Moreover, several studies used the same exact verbal and visuo-spatial task, and these two tasks correlated very inconsistently with one another across subject samples, with rs between .04 and .42 (Friedman & Miyake, 2000; Handley et al., 2002; Shah & Miyake, 1996). Such unreliable correlations obfuscate whatever the true association may be between these verbal and spatial tasks. A more subtle, but perhaps more serious, problem is that the domain-specific studies tested subject samples from a restricted range of general intellectual ability. Data were primarily collected from university students, and some from relatively prestigious universities at that. The problem with such a strategy from a psychometric perspective is that restricting the range of general ability in a sample must also restrict the contribution that general ability can make to any correlations that are observed. That is, without variation in general ability across subjects, any variability that is detected in WMC span must be due to domain-specific skills or strategies. If these same studies were conducted with more diverse subject samples, we believe that they would have yielded stronger correlations between verbal and spatial WMC measures, as well as between domain-mismatching WMC and complex-ability tests.

Our third and final reason to believe that WMC is largely domain general derived from a collection of recent studies using factor-analytic and latent-variable techniques with verbal and visuo-spatial span tasks. As a group, these studies find that latent variables comprised of verbal and visuo-spatial WMC tasks either are indistinguishable from one another, or, if separable, are very strongly correlated with one another (Ackerman, Beier & Perdue, 2002; Kyllonen, 1993; Law, Morrin & Pellegrino, 1995; Oberauer, Süß, Schulze, Wilhelm & Wittmann, 2000; Oberauer, Süß, Wilhelm & Wittmann, 2003; Park, Lautenschlager, Hedden, Davidson, Smith & Smith, 2002; Salthouse, 1995; Süß, Oberauer, Wittmann, Wilhelm & Schulze, 2002; Swanson, 1996; Wilson & Swanson, 2001). Typically, when separate verbal and visuo-spatial factors are indicated, the two share more than 65% of their variance. This is, of course, consistent with our view that both domain-general and domain-specific mechanisms are important to performance on complex span tasks of WMC, but that the lion's share of variance picked up by these tasks is quite general.

Kane et al. (2003) tested 236 subjects, from both university and community populations, in verbal and visuo-spatial tests of WMC. In contrast to many of the extant latent-variable studies of verbal versus spatial WMC, we additionally tested subjects in verbal and spatial STM tasks. These differed from the WMC tasks only in their lack

of a secondary processing demand between the presentation of each memorandum. Specifically, the verbal tasks we used were word, letter, and digit span for STM, and operation-word, reading-letter, and counting-digit span for WMC (operation-word required word memory against a equation-verification task; reading-letter required letter memory against a sentence-judgment task; counting-digit span required digit memory against an object-counting task). For the spatial domain, each STM task required subjects to reproduce sequences of visuo-spatial stimuli, such as different-sized arrows pointing in different directions, squares occupying different positions within a 4x4 matrix, and balls moving from one side of the screen to another across one of 16 paths. Each spatial WMC task presented the target memory items in alternation with a spatial processing task. The rotation-arrow task required subjects to mentally rotate letters and decide whether they were normal or mirror-reversed, and then to recall the sequence of arrows. The symmetry-matrix task required subjects to judge whether a pattern was symmetrical along its vertical axis and then recall the matrix locations. The navigation-ball task presented subjects with a version of the Brooks (1967) task, in which they saw a block letter with a star in one corner and an arrow pointing along one edge, and had to mentally navigate along the corners of the letter to report whether each corner was at the extreme top or bottom of the letter. Subjects then recalled the sequence of ball paths.

In addition to the span tasks, subjects completed a variety of standardized tasks reflecting verbal reasoning (e.g., analogies, reading comprehension, remote associates), spatial visualization (e.g., mental paper folding, mental rotation, shape assembly), and decontextualized inductive reasoning (e.g., matrix-completion tasks with novel figural stimuli, such as the Ravens Advanced Matrices). The goal was to determine whether verbal and visuo-spatial WMC differentially predicted gF, as well as reasoning in matching versus mismatching domains.

Our key predictions for the study were that: (1) verbal and visuo-spatial WMC tasks would reflect, if not a single domain-general construct, then two very strongly correlated constructs, and; (2) a latent variable derived from the domain-general WMC variance would be a strong predictor of a gF latent variable defined by the common variance among all of our reasoning tasks. Both predictions were strongly confirmed, as we detail below. We additionally explored the relation between STM, WMC and reasoning in verbal versus visuo-spatial domains. While there is clear and consistent evidence that verbal STM and WMC are distinguishable, and that WMC is the stronger predictor of general cognitive abilities (Conway et al., 2002; Engle, Tuholski et al., 1999; for a review see Daneman & Merickle, 1996), the data from spatial tasks suggest a less clear distinction between constructs. For

example, Shah and Miyake (1996) found that a spatial STM task was as good a predictor of complex spatial ability as was a spatial WMC task, and Miyake, Friedman, Rettinger, Shah and Hegarty (2001) found that spatial STM and WMC could not be distinguished at the latent variable level in a confirmatory factor analysis. Here, then, we sought to replicate these findings and begin to explore the question of why spatial STM might behave so differently from verbal STM, that is, why spatial span tasks without secondary processing demands seem to capture executive processes in ways that verbal tasks do not.

With respect to our primary question about the generality of WMC, an exploratory factor analysis conducted on all of the memory and reasoning tasks indicated that WMC reflected a single factor (comprised of the three verbal and the three spatial tasks), whereas STM was best represented by two domain-specific factors. As more rigorous tests of generality, we then conducted two series of confirmatory factor analyses on the WMC span tasks. In each series, we statistically contrasted the fit of a single-factor unitary model with the fit of a two-factor model comprised of separate verbal and spatial WMC. In the first series of analyses we allowed task-specific error to correlate when it statistically improved the fit of the model. Correlated errors reflect shared variance among pairs of tasks that is independent of the shared variance among all the tasks comprising the latent variable (recall that latent variables reflect the variance that is shared among *all* its indicator tasks). Among our verbal WMC tasks, operation span and reading span shared variance that they did not both share with counting span, perhaps because they both included word stimuli and counting span did not. Likewise, operation span and counting span shared variance that they did not share with reading span, perhaps because they both dealt with numbers. As illustrated in Figure 5 (Panel A), this first series of confirmatory factor analyses indicated that the six WMC tasks reflected a single, unitary construct rather than two. An analysis that forced the verbal and spatial WMC tasks to load onto separate factors not only failed to improve model fit, but it also yielded a correlation between the factors of .93!

In our second series of confirmatory analyses, shown in Figure 5, Panel B, we took a more conservative approach and did not allow errors to correlate. Because the correlated errors in our model were not predicted (although they were explainable post-hoc), and because the correlated errors could be interpreted as reflecting domain-specific variance (i.e., the use of words and numbers as stimuli), the inclusion of correlated errors may have biased our analyses against finding domain-specificity to improve model fit. In fact, the 2-factor model did improve fit over the 1-factor model here, indicating some domain-specificity in the WMC construct. However, the

correlation between the two factors was .84, indicating that verbal and spatial WMC shared 70% of their variance. Clearly, WMC, as measured by complex span tasks, is largely general across verbal and spatial domains. Depending on the specifics of the analyses, they may even be indistinguishable from one another.

Our second prediction was that the shared variance among WMC tasks would correlate strongly with fluid reasoning ability. This was tested in several ways. Here we did not use the two-factor WMC model that we previously found to fit the data well. This is because in structural equation modeling one cannot build interpretable models when the predictor variables are highly correlated among themselves – referred to as the multicollinearity problem. In the two-factor model, recall that verbal and spatial WMC were correlated at .84. So, our first solution to this problem was to use the domain-general WMC factor that was comprised of all six complex-span tasks (including correlated errors) to predict the gF factor derived from all of the standardized tests. This model is illustrated in Figure 6. WMC accounted for approximately 30% of the variance in gF, as in prior work (Conway et al. 2002; Engle, Tuholski et al., 1999). In addition to loading all the reasoning tasks onto a gF factor, we simultaneously loaded all the verbal tasks onto a residual, domain-specific verbal reasoning factor, representing the variance shared by the verbal tasks that was not shared by the other tasks. Similarly, we loaded all the spatial tasks onto a residual, domain-specific spatial reasoning factor, representing the variance shared by the spatial tasks that was not shared by the other tasks. Here, domain-general WMC correlated significantly with these domain-specific verbal and spatial reasoning factors (sharing ≈8% of the variance), albeit more weakly than it did with gF. We suggest that these correlations result from the contribution of WMC to learning across various domains (e.g., Daneman & Green, 1986; Hambrick & Engle, 2002; Kyllonen & Stephens, 1990).

In a subsequent test for the relations among all our memory constructs and reasoning, both WMC and STM, our solution to the multicollinearity problem was to capture the considerable shared variance among our memory tasks in a similar manner to the way we modeled our reasoning-task data, by using a nested, or "bifactor," structure. Nested models allow tasks to simultaneously load onto more than one factor, and so variance attributable to different underlying constructs can be extracted independently from each task. The logic of our analysis was that no WMC or STM task provides a pure measure of either domain-general executive attention or domain-specific storage and rehearsal; all memory-span tasks will reflect storage, rehearsal, and executive attention to some degree (indeed, all cognitive tasks may reflect executive attention to some degree). By our view, WMC tasks capture executive attention primarily but also domain-specific rehearsal and storage,

whereas STM tasks capture domain-specific storage primarily but also executive attention. As illustrated in Figure 7, our nested model thus consisted of an "Exec-Attn" factor, with loadings from *all* memory variables, reflecting the domain-general "executive" variance shared by all the STM and WM tasks. The model also consisted of domain-specific storage/rehearsal factors, with loadings from the six verbal span tasks on the "Storage-V" factor and loadings from the six spatial span tasks on the "Storage-S" factor. Thus, from each task we extracted variance hypothesized to reflect domain-general executive-attention *and* variance hypothesized to reflect storage, rehearsal, or coding processes that were specific to either verbal or spatial stimuli. The Exec-Attn factor yielded high factor loadings from verbal and spatial WMC tasks and low loadings from verbal and spatial STM tasks, indicating empirically that it represented primarily domain-general attention control. In contrast, the domain-specific storage factors each elicited high loadings from their respective STM tasks and lower loadings from their WMC tasks, indicating that they reflected primarily domain-specific storage and rehearsal processes.

As illustrated in Figure 7, the executive factor correlated substantially with gF (≈30% shared variance) and significantly, but more weakly, with domain-specific reasoning (≈8% shared variance). Thus, this executive-attention factor behaved very similarly to our unitary WMC factor from our previous analysis. These two models together clearly indicate that the domain-general executive processes shared among WMC tasks, and not the domain-specific storage and rehearsal processes they also measure, are what drives the correlation between WMC and general fluid intelligence.

Another interesting feature of this structural model is that the verbal and spatial storage factors showed quite divergent patterns of correlations with reasoning. Verbal storage predicted unique variance in verbal reasoning beyond that accounted for by WMC, but it did not significantly predict unique variance in gF. Both findings are consistent with our prior work (Cantor, Engle & Hamilton, 1991; Engle et al., 1990; Engle, Tuholski et al., 1999). In contrast, spatial storage not only predicted unique variance in spatial reasoning, it also accounted for as much unique variance in gF as did executive attention. The variance associated with spatial storage appears to be quite general in its predictive power, correlating with both domain-specific and domain-general aspects of complex reasoning (see also Miyake et al., 2001; Oberauer, 1993; Shah & Miyake, 1996).

How can we account for the apparent generality of spatial storage? Why do these "simple" span tasks work so well in predicting complex cognition? Shah and Miyake (1996) argued that subjects who do well on spatial STM tasks may be more strategic than are those who do poorly, perhaps employing spatial chunking or

some other beneficial coding processes, and this strategic superiority also improves performance in complex ability tasks. Another possibility is that spatial STM measures are purer measures of executive attention than are verbal STM measures. That is, spatial STM tasks with abstract, novel stimuli do not benefit from either the well-learned rehearsal strategies that are available to verbal materials (such as inner speech, associative chaining, etc.), nor do they afford the use of semantic or lexical knowledge to help encode or retrieve list items. Spatial tasks therefore may rely more on "brute force" executive-fueled maintenance than on specialized rehearsal routines. By this view, spatial STM is really an executive task similar to WMC tasks. We find this to be an attractive view, and one that is consistent with the spatial WMC/STM findings of Miyake et al. (2001). The difficulty with it, however, is that in our data, as in Shah and Miyake (1996), spatial storage accounts for different variance in gF than does executive attention/WMC. If spatial storage was simply another executive-attention measure, then it should account for much of the same gF variance that WMC tasks do.

A very different solution to these questions about spatial STM, at least for our data, is that our gF factor may have been more biased to the spatial domain than to the verbal domain. If true, then what looked like "general" reasoning ability being predicted by spatial storage was, instead, largely spatial reasoning. Although our gF factor consisted of five putatively verbal and five putatively spatial reasoning tasks, one of the verbal tasks (syllogisms) loaded with the spatial tasks in our exploratory factor analysis. Plus, the three matrix-reasoning tasks that loaded onto gF also consisted of some items that involved visuo-spatial processing (this was especially true of the Ravens test). We therefore used our nested model of memory span, consisting of executive attention, spatial storage, and verbal storage, to predict gF factors derived from different combinations of reasoning tasks.

In the first model, the gF latent variable was extracted from the three matrix reasoning tasks, which are "gold standard" gF tasks that nonetheless may have some spatial component. Here, the correlations of gF with executive attention, spatial storage, and verbal storage were .55, .54, and .17, respectively; spatial storage accounted for as much unique variance in gF as did executive attention. In the second model, however, we balanced the verbal/spatial contribution to gF by extracting it from three verbal and visuo-spatial measures; no matrix tasks were used. Here, the resulting correlations with memory factors were .57, .47, and .24, respectively. Although spatial storage still accounted for substantial variance in this more balanced gF factor, its contribution was reduced relative to model 1 and relative to the executive-attention contribution. Note that the executive-attention contribution did not change between analyses. In a third and final model, we defined gF using the three

verbal tasks from model 2, in addition to the remote associates task, a putatively "verbal" task that nonetheless measured domain-general inductive reasoning according to our exploratory factor analysis. The correlations with this more verbal gF factor were .51, .29, and .36, respectively. Clearly, spatial storage still does share variance with fluid verbal abilities, but it accounts for less and less gF variance as gF became more verbal (with correlations of .54, .47, and .29). In contrast, the executive attention factor shared 25 – 30% of the variance in gF (with correlations of .55, .57, and .51) no matter how gF was defined. These analyses suggest that spatial storage may be a bit more general in its predictive power than is verbal storage, but it is not as general as the executive-attention contribution to memory span.

Altogether then, the Kane et al. (2003) data strongly indicate that verbal and visuo-spatial WMC tasks share a core, domain-general set of processes that represent more than simple STM storage and rehearsal. We would argue that the shared variance among WMC tasks represents domain-general executive attention, which is an important determinant of general fluid intelligence and reasoning ability. Although the contributions of verbal and spatial storage to memory span also predict variance to reasoning ability, these correlations are stronger with reasoning in the matching stimulus domain than with domain-general thinking abilities. Spatial storage does appear to be somewhat "special" in its relation to general ability, but our final set of analyses indicates spatial storage to be less general in predicting complex cognition than is the executive-attention contribution to memory span.

IV. Microanalytic Studies of Working Memory Capacity: Its Relation to Executive Attentional Control

We have argued, based on our large-scale macroanalytic studies, that the critical element of complex WMC span tasks for higher-order cognition and general fluid abilities, whether spatial or verbal, is the domain-general ability to control attention. That conclusion was inferential at the time we proposed it (Engle, Kane et al., 1999; Engle, Tuholski et al., 1999), but we had no direct evidence for support. There is now considerable data to support that thesis and we will describe it here.

A. WMC and Retrieval Interference

As we discussed at length in our introduction to this chapter, it is now clear that WMC is an important factor in the degree to which an individual's recall performance will be diminished by proactive interference. One line of research supporting this conclusion is based on "fan effect" manipulations (Anderson, 1983), whereby cues that are associated with many items or events allow slower recognition than do cues associated with few items or

events. Bunting et al. (2003) and Cantor and Engle (1993) both showed that low WMC subjects show a much steeper fan effect than high span subjects for propositional information if there is overlap among the fan items in set membership. However, if all the items are unique to a given fan, thereby eliminating response competition between sets, then high and low spans do not differ.

Conway and Engle (1994) demonstrated the importance of competition, or conflict, to eliciting WMC differences in fan effects by having subjects learn to associate letters with a digit cue representing the number of items in a set. Thus, *C* and *S* might be associated with the digit 2, *W*, *G*, *H*, and *X* with 4, and so on. After an extensive learning phase, subjects saw a digit (e.g., 2) and a letter (e.g., C) and they were to press a key indicating whether or not the letter was in the set represented by the digit. When there was no overlap among the set items, i.e., a letter was unique to a given set, the set size function for high and low WMC subjects did not differ. Moreover, the performance of high span subjects was not further disrupted in a condition with conflict, in which each item was a member of two different sets. However, the set size function for low spans was substantially steeper in the response competition condition – they showed greater interference than did high spans, and they showed greater interference than they did under no competition. In other words, high and low span subjects showed similar search rates of active memory in the absence of interference, but low spans were differentially slowed under conditions of interference, or what we might think of as response competition. Conway and Engle argued that high spans were able to attentionally inhibit the conflict from competing sets in the overlap condition, but low spans were not, and so low spans were more vulnerable to blocking and/or confusion among competing sets.

Kane and Engle (2000) provided a more direct demonstration of the role of attention control in the interaction between WMC and interference vulnerability. Our subjects read a 10-word list from a category such as "animals," then performed a 15 s rehearsal-preventative task, and then were cued to recall the 10 words. They received a series of such lists, all drawn from the same category, thereby inducing proactive interference across lists. On the very first such list, both high and low span subjects recalled approximately 6 words – not different from one another and not near ceiling or floor. On subsequent lists, the recall by low spans fell off at a faster rate than that of high spans. In other words, low spans showed a steeper interference function than high spans.

Some of our subjects additionally performed an attention-demanding secondary task either during the encoding or retrieval phase of the memory task. The interference function, i.e. the change from trial 1 to trial 2 to

trial 3, did not change for low spans under attentional load compared to low spans under no load. However, the load manipulation caused the interference function for the high spans to become considerably steeper and virtually identical to that of the low spans. Thus, under standard conditions low spans were more vulnerable to interference than were high spans, but under load, the span groups were equivalently vulnerable. Our interpretation of these findings was that, in the absence of an attention-demanding secondary task, high WMC individuals were capable of controlling their attention in such a manner that they encoded new list items as distinct from earlier list items and, during retrieval, blocked intrusions from the interfering lists. However, under load, high spans were incapable of using control in these ways. We further argued that low span subjects were less capable of engaging attentional processes to resist interference, and so by failing to use controlled processing under normal conditions they were not able to be hurt further by the load of the secondary task. Interestingly, low spans showed a larger dual-task decrement than high spans on list 1 of the task, before interference had built up. This suggests that low spans may have been exhausting their attention-control capabilities simply to encode and retrieve a single list of associated items, even in the absence of interference, and so they essentially had nothing left to give to combat the added effects of interference on subsequent lists.

The Kane and Engle (2000) finding that low spans have more difficulty than high spans in blocking the effects of prior-list information is consistent with previous findings reported in two papers by Rosen and Engle. In the first (1997), they conducted a series of studies using a fluency retrieval task. Subjects were to recall as many different exemplars of the category "animals" as possible in 10 minutes, with instructions to not repeat any items. In three experiments, high span subjects retrieved many more animals than did low spans. In a fourth experiment, subjects were instructed that, while we were interested in how many different animals they could name, if an already recalled item came to mind, they should say it anyway "to clear their minds." High spans made relatively few re-retrievals but low spans repeated nearly half their retrieved items. Again, we reasoned that high spans had sufficient attentional resources to monitor for previously retrieved items and to suppress their activation. However, low spans did not have sufficient attentional resources to both monitor for whether a retrieved item had been previously retrieved and also to suppress activation of those items. This series of studies also found that, while a secondary-load task greatly reduced the number of exemplars retrieved by high spans, it had little effect on retrieval by low spans. This suggested, as in the Kane and Engle (2000) study, that high spans were using their ability to focus and maintain attention for controlled strategic retrieval as well as for suppression

of previously retrieved items. Low spans were not using such attention control to strategic retrieval or suppression during the regular version of the task, and so their performance was not impaired further by divided attention.

Traditional paired-associates tasks also support the conclusion that low span subjects are impaired in the attentional blocking of competition during memory retrieval. Rosen and Engle (1998) had subjects learn three lists of paired associates using an A-B, A-C, A-B design with responses given orally in response to the cue word and, in the first experiment, a response deadline of 1300 msec. List 1 was composed of items with high pre-experimental associations, e.g., "*bird-bath*" and "*knee-bend*." High and low span subjects did not differ in the trials to learn this first list. The second list consisted of the cue words from list one associated with new words that were weak associates, e.g., "*bird-dawn*" and "*knee-bone*." The interference from list 1 caused both groups to take longer to learn list 2, but low spans took substantially longer to learn than high spans, indicating a relation between WMC and negative transfer (or proactive interference at learning). Further, low spans made many more intrusions from list 1 during the learning of list 2 than did high spans. The third list consisted of re-learning the items from list one (*bird-bath, knee-bend*). Even though both groups had previously learned this list and in an equivalent number of trials, low spans now required more trials to re-learn the list and, in so doing, made more intrusions than did high spans.

B. WMC and Inhibition/Suppression

The notion of inhibition or deactivation of a representation remains a controversial topic in cognitive research (MacLeod, Dodd, Sheard, Wilson & Bibi, in press). For example, in learning the second list of the Rosen and Engle (1998) study described above, high spans could make few intrusions of "*bath*" to "*bird*" because they have dampened that connection (Postman, Stark & Frasier, 1968). Or, they could make few intrusions, instead, because they quickly strengthen the "*bird-dawn*" connection to a higher level. Most techniques for studying so-called inhibition do not allow a distinction between a mechanism based on true inhibition and one based on an increase in excitation. We have taken the position that both mechanisms require the control of attention and therefore will depend on WMC.

We have shown that the negative priming effect, in which a distractor letter to be ignored on trial n is the target letter to be named on trial $n+1$, is resource-dependent (Engle et al, 1995); that is, the effect disappears under a secondary load task. Further, whether subjects show the negative priming effect depends on their WMC,

with high span subjects showing larger effects than low spans (Conway, Tuholski, Shisler & Engle (1999). Perhaps the strongest evidence for what appear to be true inhibition differences is the second study of the Rosen and Engle (1998) paper. We used an identical A-B, A-C, A-B procedure to that described above, except that instead of forcing the subjects to respond quickly so that we could focus on intrusions, we emphasized accuracy of response so that we could measure time to retrieve the item. If high spans suppress activation of the "*bird-bath*" connection from list 1 during the learning of "*bird-dawn*" in list 2, then, when we test them on list 3, which is the relearning of list 1, they should be slower than a control group of high spans learning the "*bird-bath*" connection for the first time. They may even be slower to respond than they themselves had been on the first recall phase of list 1. In contrast, if low WMC individuals have less capability to suppress the list 1 items during the learning of list 2, then they might show less of an increase in the time to retrieve list 1 items in the first recall phase of list 3 learning. That is exactly what Rosen and Engle found. Low spans in the interference condition were actually faster than in the non-interference condition to retrieve "*bath*" as a response to "*bird*" on list 3. However, high spans in the interference condition were significantly slower to retrieve list 3 responses during the first recall phase than the non-interference group. In addition, high spans in the interference condition were slower to retrieve "*bath*" to "*bird*" during the first recall phase of list 3 learning than they were themselves during the first recall phase of learning the same items on list 1. This strikes us as strong evidence that high spans suppressed the list 1 ("*bird-bath*") connections during the learning of list 2 and that low spans learned the A-C list with relatively little evidence of suppression of the A-B connection.

C. WMC and Resistance to Prepotent Responses

If our thesis that performance on complex WMC tasks such as operation span and reading span reflect primarily an ability to control attention, irrespective of mode of representation, then we should find that high and low spans perform differently on tasks that require responses counter to strongly established stimulus-response connections. That is, WMC differences should be measurable in "attention control" tasks that are further removed from a memory context. We will describe our work using the antisaccade task and the Stroop task to support this contention.

The antisaccade task is perhaps the best possible task with which to test this idea. Millions of years of evolution have prepared us to attend to any stimulus that cues movement. After all, moving objects might be predator or prey, and so survival depends attending to them. The task is as follows: You are seated in front of a

computer monitor and asked to look at a fixation point. At some time, there is a flickering cue, 11° to one side or the other, randomly. Your natural tendency is to shift your attention and to move your eyes to the flickering cue. However, your task is instead to immediately move your eyes to the *opposite* side of the screen, thus disobeying Mother Nature's instructions. The antisaccade task typically has two conditions: the prosaccade condition, in which you are to move your eyes to the flickering cue, and the antisaccade condition, in which you are to shift your attention and eye gaze to the opposite side of the screen.

If WMC reflects individual differences in ability to control attention, then people who score high on a complex WMC span task should perform better on the antisaccade task than do those who score low in complex span. At the same time, high and low spans should not differ on the prosaccade task, because here attention can be drawn or captured by the exogenous event, resulting in the automatic fixation at the location of the target. Kane et al (2001) used a procedure in which one of three visually similar letters, B, P, or R, was presented either at the location of the previous flickering cue (prosaccade condition) or at the equivalent location on the opposite side of the screen (antisaccade condition). The letter occurred very briefly and was pattern-masked, so if the subject shifted attention toward the exogenous cue even briefly while in the antisaccade condition, they would likely misidentify the letter or at least have a slowed response. We found that the two groups were not different in the prosaccade condition, either in number of errors or in time to initiate correct responses. However, in the antisaccade condition, low spans made more identification errors and were slower on correct trials than did high spans. Nearly an hour of antisaccade practice still showed that high spans made fewer reflexive saccades to the flickering cue than low spans. And, even on trials in which both high and low spans made an accurate antisaccade, high spans did so significantly more quickly than did low spans.

One potential problem with the Kane et al (2001) procedure is the possibility that low spans had more difficulty than high spans with the letter-identification task. Roberts, Hager and Heron (1994) demonstrated that subjects under a secondary, attention-demanding load made more antisaccade errors than did subjects under normal conditions. Therefore, if the letter task was more demanding for the low spans than for the high spans, this could have resulted in low spans making more antisaccade errors. To correct for this potential problem, Unsworth, Schrock and Engle (2003) developed a task in which subjects simply had to move their eyes to a box located 11° left or right of fixation. In the prosaccade condition, subjects were to move their fixation as quickly as possible to the box that flickered. In the antisaccade condition, subjects were to move their gaze to the box on

the opposite side of the screen from the box that flickered. Figure 8 shows the percentage of errors on the first saccade. Consistent with our prior work, high and low spans did not differ in the prosaccade condition. They were equivalent in the accuracy of direction of the first saccade and in the time to initiate that first saccade. In the antisaccade condition, however, low spans made many more errors in their first saccade. In addition, even if the first saccade was in the correct direction, low spans were slower to initiate that saccade. These findings are consistent with those of Kane et al. (2001), and suggest that the span differences we originally found were not an artifact of the embedded letter-identification task.

V. A Two-Factor Theory of Executive Control

Our antisaccade findings also support a two-factor model of the executive control of behavior, which also seems to explain the Stroop results we will describe below. We propose one factor of control to be the maintenance of the task goals in active memory, and that low span subjects are simply less able to maintain the novel production necessary to do the task ("*Look away from the flash*") in active memory. All subjects clearly knew what they were supposed to do in the task, and they could easily tell you what they were to do, presumably based on retrieval of the goal from LTM. However, in the context of doing the antisaccade task, trial after trial, low spans failed on some trials to do the mental work necessary to maintain the production in active memory such that it could control behavior. Under these circumstances low span subjects were more likely to make a saccade to the cue, in error, than were high spans. Our view is that maintenance is a resource-demanding endeavor and that high WMC individuals are better able to expend that resource on a continuing basis. We believe that the prefrontal cortex is important in successful maintenance of the task goals in active memory and will have more to say about that below.

The second factor in the executive control of behavior is the resolution of response competition or conflict, particularly when prepotent or habitual behaviors conflict with behaviors appropriate to the current task goal. We argue that, even when the production necessary to perform the antisaccade task is in active memory, there is conflict between the natural, prepotent response tendency to attend to and look toward the flickering exogenous cue and the response tendency resulting from the task goal provided by the experimental context. Low spans have greater difficulty resolving that conflict as demonstrated by the fact that even when they made the correct initial saccade, indicating effective goal maintenance, they were slower to initiate the saccade than were high spans.

Our studies with the Stroop (1935) paradigm show a striking parallel to our studies using the antisaccade task, and in fact they were explicitly designed to test our dual-process idea. Kane and Engle (2003) tested high and low span subjects in different versions of the Stroop color-word task, in which subjects name the colors in which words are presented (e.g., *RED* presented in the color blue). These studies were motivated, in part, by failures in the psychometric and neuropsychological literatures to demonstrate a consistent relationship between Stroop performance and either intelligence or prefrontal cortex damage. These failures were interesting to us – and initially surprising – because both intelligence and prefrontal cortex have been strongly implicated in WMC and attention-control functions (for a review, see Kane & Engle, 2002). However, our reading of the relevant literatures suggested to us that studies that found no relation between Stroop performance and intelligence or prefrontal function tended to use versions of the Stroop task in which all (or almost all) of the words and colors were in conflict. We thought that this was significant because, by our view, part of the challenge in the Stroop task is to actively maintain a novel goal ("*name the color*") in the face of a powerful opposing habit (i.e., to read the word). Therefore, a task context in which all the stimuli reinforced the task goal by presenting only incongruent stimuli would minimize the need for active goal maintenance. When trial after trial forces subjects to ignore the word, ignore the word, and again, ignore the word, the task goal may become overlearned and thus run off without active, controlled maintenance.

Consider, in contrast, a task context in which a majority of trials are _congruent_, with the word and color matching each other (e.g., *BLUE* presented in blue). Here a subject could respond accurately on most trials even if they completely failed to act according to the goal, and instead slipped into reading the words rather than naming the colors. When that subject encountered one of the rare incongruent stimuli, it is unlikely that he or she could respond both quickly and accurately. For a subject to respond quickly and accurately to an infrequent incongruent stimulus in a high-congruency task, he or she must actively maintain accessibility to the goal of the task. Otherwise, the habitual and incorrect response will be elicited. We therefore predicted that, as in the antisaccade task, low span subjects would show evidence of failed goal maintenance in the Stroop task, but perhaps only in a high-congruency context. We expected that when most Stroop trials were congruent, low spans would make many more errors on incongruent trials than would high spans. Moreover, by the dual-process view of executive control, even in contexts in which goal maintenance was less critical, for example in a low-

congruency context, a span difference in resolving response conflict might be evident in response-time interference.

In fact, this is exactly what Kane and Engle (2003) observed. In task contexts where 75% or 80% of the trials were congruent, low spans showed significantly greater interference, as measured by errors, than did high spans. The results from four such conditions (each with different groups of subjects) are presented in Figure 9. Moreover, in one experiment we had a large enough subject sample to examine the latencies of errors in the 80%-congruent condition, with the expectation that errors resulting from goal neglect (and subsequent word reading) should be relatively fast compared to other kinds of errors. We therefore expected that when subjects' errors represented unambiguous, "clean" responses of reading the word on incongruent trials, they would be faster than other errors such as stuttering, slurring two words together, or naming a word that was not presented. We also predicted that low spans would show more of these "clean" errors than would high spans. To test this idea, we examined error latencies for subjects who made at least a 16% error rate on incongruent trials. Twenty-two high spans and 47 low spans met this criterion, and on average, 68% of low span subjects' errors, but only 58% of high spans' errors, were "clean", or indicative of goal-maintenance failure. Irrespective of WMC, clean errors were committed over 1000 ms faster than were other errors, and with latencies very similar to correct responses on congruent trials, strongly suggesting that these errors represented rapid word reading due to failed access to the goal state.

As a final source of evidence for failed goal maintenance, low spans also demonstrated greater response-time facilitation than did high spans in the high-congruency conditions. That is, low spans showed a differential latency benefit on congruent trials, where word and color match, compared to neutral trials. What does facilitation have to do with goal maintenance? MacLeod (1998; MacLeod & MacDonald, 2000) has argued that facilitation in the Stroop task reflects the fact that subjects sometimes read the word on congruent trials rather than naming the color, and because word reading is faster than color naming, these undetectable reading responses reduce the mean latency for congruent trials. Put into our words, the word reading responsible for facilitation effects is a result of periodic failure of goal maintenance. Low spans should therefore show greater facilitation than do high spans and that is just what we found. Moreover, collapsed across span groups, we found significant correlations between error interference and response-time facilitation in our high-congruency conditions (rs between .35 and .45), the two measures we hypothesized to reflect word reading due to failures of goal maintenance.

We also found evidence for span differences in resolving response competition under conditions where goal-maintenance failures were unlikely, supporting our idea that WMC is related to two aspects of executive control. In Stroop contexts that reinforced the task goal by presenting 0% congruent trials, we found modest span differences in response-time interference. These differences were on the order of only 20 – 30 ms, and they required much larger samples to be statistically significant than did the error effects we discussed previously. Our idea is that these low-congruency contexts did not put a premium on actively maintaining access to the task goals, and so the latency differences we observed between high and low spans reflect low spans' deficiency in resolving response competition (as in our antisaccade and memory-interference studies). Further support for this idea came from two experiments in which a 75% congruent context was presented to subjects after they had extensively practiced a 0% congruent Stroop task. Here, overlearning of the task goal in the prior context might make goal maintenance in the 75% congruent condition less necessary. And, in fact, low spans and high spans showed equivalent (and low) error rates in the 75% conditions here, in addition to showing equivalent response-time facilitation effects. High and low spans did differ, however, in response-time interference, suggesting to us that low spans were responding according to goal, but they were slower to resolve the competition between color and word than were high spans.

Our Stroop and antisaccade findings generally indicate that high and low WMC subjects differ not only in higher order, complex cognitive tasks, but also in relatively "lower order," simple attention tasks. Specifically, when powerful habits, prepotencies, or reflexes must be held in abeyance in order to satisfy current goals, high spans more effectively exert executive control than do low spans. Moreover, our view is that such executive control reflects a synergy of "memorial" and "attentional" processes. Active maintenance of goals, a memory phenomenon, allows the resolution of response competition to occur – without effective goal maintenance, automated routines will control behavior in the face of conflict. However, even when goal maintenance is successful, the attentional implementation of blocking or inhibitory processes may sometimes fail, or at least they may be slow to resolve the competition that is present. It is our view that both of these control processes rely on WMC.

VI. Implementation of Working Memory Capacity in the Brain

We have so far discussed our dual-process view of executive control as if it was entirely new, but this is really not the case. The behavioral and neuroscience research programs of both John Duncan and Jonathan

Cohen have heavily influenced our thinking about WMC and executive attention, at least insofar as they relate to the idea of goal maintenance. These views also provide suggestions for how our ideas might be mechanistically implemented in the wetware of the brain. Duncan (1993, 1995) has argued that in novel contexts, or in those that afford multiple actions, attention-control processes somehow weight a hierarchical organization of goal abstractions, and this weighting serves to bias the system toward goal attainment. Important to our perspective, Duncan argues that such attentional, controlled goal weighting is strongly associated with general fluid intelligence and relies heavily on prefrontal cortex circuitry. Evidence for Duncan's ideas come from studies showing that dual-task conditions, low fluid intelligence, and prefrontal cortex damage lead to high rates of "goal neglect" in novel tasks, even when subjects can faithfully report what the goal of the task actually is (probably based on LTM retrieval; Duncan, Burgess & Emslie, 1995; Duncan, Emslie, Williams, Johnson & Freer, 1996). By our view that WMC, attention control, fluid intelligence, and prefrontal cortex functioning are largely overlapping constructs (Engle, Kane et al., 1999; Engle & Oransky, 1999; Kane & Engle, 2002), this confluence of influences on goal neglect indicate the centrality of WMC to goal maintenance, and the importance of such maintenance for complex, intentional behavior.

Cohen's research on the Stroop task and on the cognitive neuroscience of executive control also suggests a link between goal maintenance and prefrontal cortex functioning. In essence, Cohen's connectionist models and imaging research suggest that the dorsolateral area of the prefrontal cortex is particularly involved in the on-line maintenance of "task demand", or contextual information that keeps behavior yoked to goals (Braver & Cohen,. 2000; Cohen & Servan-Schreiber, 1992; O'Reilly, Braver & Cohen, 1999). For example, Cohen models the Stroop deficits seen in schizophrenics by reducing the activity of task-demand context nodes ("*name the color*"). This reduction in activity represents in the model schizophrenics' decreased dopaminergic activity in prefrontal cortex circuitry. When these task-demand nodes operate effectively, in a healthy brain, they block activity of pathways associated with the environmentally elicited, but incorrect, response. When "damaged" by schizophrenia, prefrontal cortex damage, or presumably, low WMC, however, these task-demand representations of goal states can no longer block the dominant, prepotent response, leading to exaggerated Stroop interference effects. Mechanistically, then, the executive control of behavior is implemented via the active maintenance of goals (Braver & Cohen, 2000; O'Reilly, Braver & Cohen, 1999).

A particularly compelling empirical confirmation of Cohen's ideas was reported recently by MacDonald, Cohen, Stenger & Carter (2000). Under fMRI, subjects completed a 50%-congruent Stroop task in which the instructions to read the word or name the color were presented 11 s before each stimulus. On color-naming trials, where active goal maintenance would seem most necessary, prefrontal cortex activity increased steadily over the 11 s delay. On the more automatic word-reading trials, however, no such increase in activity was observed. Thus, prefrontal cortex activity seems to have reflected a mounting preparation to respond according to the novel goal to "*name the color, not the word.*" This interpretation is bolstered by the additional finding that delay-period prefrontal activity was negatively correlated with Stroop interference ($r = -.63$). That is, the more active prefrontal cortex was before the Stroop stimulus arrived, the less Stroop interference was elicited. Related findings have been reported by West and Alain (2000), who used event-related potentials to isolate a slow wave originating in prefrontal cortex that predicts, in advance, when a Stroop error is about to be committed. Specifically, this wave begins 400 – 800 ms before the error-eliciting stimulus is presented, and it is significantly larger in high-congruency than in low-congruency Stroop tasks. Given our findings of WMC differences in error interference under high-congruency conditions, the imaging findings discussed here strongly suggest that WMC differences in executive control are linked to individual differences in prefrontal cortex activity corresponding to active goal maintenance.

The second component of our theory involves differences in the resolution of conflict, evident in antisaccade and Stroop tasks as slower responding for low spans when faced with competition, even when they appear to have acted according to goal. Our interpretation of the memory interference and retrieval inhibition findings that we discussed above also would suggest response competition or conflict as the likely culprit responsible for the differences between high and low WMC subjects. For example, in the Rosen and Engle (1998) interference study, once a person has learned to give "*bath*" in response to "*bird*," then during the period when the subject must learn to give "*dawn*" to "*bird*," we believe that high and low WMC subjects differ in their ability to detect and resolve the conflict arising from the retrieval of "*bath*" to "*bird*." High spans appear to be able to suppress the inappropriate retrieval better than the lows.

The detection and resolution of conflict appears to rely on anterior cingulate, as also indicated by recent work from Jonathan Cohen's group (Botvinick, Braver, Barch, Carter, and Cohen, 2001; see also MacLeod & MacDonald, 2000). They also reported two computational modeling studies supporting that view. The argument

is that the anterior cingulate detects overall conflict in the system and, through a feedback loop, causes increased activity in other regions, such as the prefrontal cortex. That, in turn, would lead to better maintenance of novel connections, task goals, and productions. This neural interaction of competition detection/resolution and goal maintenance seems a likely mechanism by which individual differences of the kinds we have described here could be implemented in the nervous system.

VII. Conclusion

Measures of short-term memory such as digit and word span correlate very poorly with real-world cognitive tasks but measures of working memory capacity correlate with a wide array of such tasks. Measures of WMC are highly reliable and highly valid indicators of some construct of clear relevance to feral cognition. Our macroanalytic studies have demonstrated that the construct reflected by WMC tasks has a strong relationship with general fluid intelligence above and beyond what these tasks share with simple span tasks. Further, this construct is domain-free and general and is common to complex span tasks both verbal and spatial in nature. Our microanalytic studies provide evidence that the construct reflects the ability to control attention, particularly when other elements of the internal and external environment are serving to capture attention away from the currently-relevant task. We have referred to this as executive attention and think of it as the ability to maintain stimulus and response elements in active memory, particularly in the presence of events that would capture attention away from that enterprise. We proposed a two-factor model by which individual differences in WMC or executive attention leads to performance differences. We argued that executive attention is important for maintaining information in active memory and secondly is important in the resolution of conflict resulting from competition between task-appropriate responses and prepotent but inappropriate responses. The conflict might also arise from stimulus representations of competing strength. This two-factor model fits with current thinking about the role of two brain structures: the prefrontal cortex as important to the maintenance of information in an active and easily accessible state and the anterior cingulate as important to the detection and resolution of conflict.

References

Ackerman, P. L, Beier, M. E., & Perdue, M. B. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. Journal of Experimental Psychology: General, 131, 567-589.

Allport, D. A., Styles, E.A., & Hsieh, S. (1994). Shifting attentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), Attention and performance XV (pp. 421-452). Hillsdale, NJ: Erlbaum.

Arnett, P.A., Higgenson, C.I., Voss, W.D., Bender, W.I., Wurst, J.M., & Tippin, J.M. (1999). Depression in multiple sclerosis: Relationship to working memory capacity. Neuropsychology, 13, 546-556.

Anderson, J. R. (1983). The Architecture of Cognition. Cambridge, MA: Harvard University Press.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), The psychology of learning and motivation (vol. 2, pp. 89-95). New York: Academic Press.

Baddeley, A. D. (1986). Working memory. London/New York: Oxford University Press.

Baddeley, A. D. (1996). Exploring the central executive. Quarterly Journal of Experimental Psychology, 49A, 5-28.

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? Trends in Cognitive Sciences, 4, 417-423.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. A. Bower (Ed.), The psychology of learning and motivation (vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A. D., & Hitch, G. (1977). Recency re-examined. In S. Dornic (Ed.) Attention and Performance VI (pp. 647-667). Hillsdale, NJ: Erlbaum.

Baddeley, A. D., & Logie, R. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control (pp. 28-61). New York: Cambridge University Press.

Baddeley, A. D., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. Psychologial Review, 105, 158-173.

Barch, D.M., Sheline, Y.I., Csernansky, J.G., & Snyder, A.Z. (2003). Working memory and prefrontal cortex dysfunction: Specificity to schizophrenia compared with major depression. Biological Psychiatry, 53, 376-384.

Barrouillet, P. (1996). Transitive inferences from set-inclusion relations and working memory. Journal of Experimental Psychology: Learning, Memory, and Cognition, 22, 1408-1422.

Benton, S. L., Kraft, R. G., Glover, J. A., & Plake, B. S. (1984). Cognitive capacity differences among writers. Journal of Educational Psychology, 76, 820-834.

Botvinick, M.M.; Braver, T.S.; Barch, D.M.; Carter, C. S.; Cohen, J. D. (2001). Conflict monitoring and cognitive control. Psychological Review, 108, 624-652.

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dompamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), Attention and Performance XVIII: Control of Cognitive Processes (pp. 713-737). Cambridge, MA: MIT Press.

Brewin, C. R., & Beaton, A. (2002). Thought suppression, intelligence, and working memory capacity. Behavior Research and Therapy, 40, 923-930.

Bunting, M. F., & Conway, A. R. A., & Heitz, R. P. (2003). Individual differences in the fan effect and working memory capacity. Manuscript in preparation.

Cantor, J., & Engle, R. W. (1993). Working-memory capacity as long-term memory activation: An individual-differences approach. Journal of Experimental Psychology, Learning, Memory, and Cognition, 19, 1101-1114.

Cantor, J., Engle, R. W., & Hamilton, G. (1991). Short-term memory, working memory, and verbal abilities: How do they relate? Intelligence, 15, 229-246.

Caplan, D., & Waters, G. S. (1999). Verbal working memory and sentence comprehension. Behavioral and Brain Sciences, 22, 77-126.

Carpenter, P. A., & Just, M. A. (1988). The role of working memory in language comprehension. In D. Klahr & K. Kotovsky (Eds.), Complex information processing: The impact of Herbert A. Simon (pp. 31-68). Hillsdale, NJ: Erlbaum.

Case, R. (1985). Intellectual development: Birth to adulthood. New York: Academic Press.

Clark, C.R., McFarlane, A.C., Morris, P., Weber, D.L., Sonkkilla, C., Shaw, M., Marcina, J., Tochon-Danguy, H.J., & Egan, G.F. (2003). Cerebral function in posttraumatic stress disorder during verbal working memory updating: A positron emission tomography study. Biological Psychiatry, 53, 474-481.

Clarkson-Smith, L., & Hartley, A. A. (1990). The game of bridge as an exercise in working memory and reasoning. Journal of Gerontology, 45, P233-P238.

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. Journal of Experimental Psychology: General, 123, 354-373.

Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. Psychonomic Bulletin and Review, 8, 331-335.

Conway, A. R. A., Cowan, N., Bunting, M. F., Therriault, D., & Minkoff, S. (2002). A latent variable analysis of working memory capacity, short term memory capacity, processing speed, and general fluid intelligence. Intelligence, 30, 163-183.

Conway, A. R. A., Tuholski, S. W., Shisler, R. J., & Engle, R. W. (1999). The effect of memory load on negative priming: An individual differences investigation. Memory & Cognition, 27, 1042-1050.

Conway, A.R.A., Kane, M.J., & Engle, R.W. (1999). Is Spearman's g determined by speed or working memory capacity? Book review of Jensen on Intelligence-g-Factor. Psycoloquy, 10 (074), article 16.

Cowan, N. (1995). Attention and Memory: An Integrated Framework. Oxford: Oxford University Press.

Cowan, N. (1999). An embedded-process model of working memory. In A. Miyake & P. Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control (pp. 62-101). New York: Cambridge University Press.

Craik, F. I. M., & Watkins, M. J. (1973). The role of rehearsal in short-term memory. Journal of Verbal Learning and Verbal Behavior, 12, 599-607.

Cronbach, L. J., (1957). The two disciplines of scientific psychology. American Psychologist, 12, 671-684.

Crowder, R. G. (1982). The demise of short-term memory. Acta Psychologica, 50, 291-323.

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. Journal of Verbal Learning and Verbal Behavior, 19, 450-466.

Daneman, M., & Carpenter, P. A. (1983). Individual differences in integrating information between and within sentences. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 561-584.

Daneman, M., & Green, I. (1986). Individual differences in comprehending and producing words in context. Journal of Memory and Language, 25, 1-18.

Daneman, M., & Merikle, P. M. (1996). Working memory and language comprehension: A meta-analysis. Psychonomic Bulletin & Review, 3, 422-433.

Daneman, M., & Tardif, T. (1987). Working memory and reading skill reexamined. In M. Coltheart (Ed.), Attention and performance XII: The psychology of reading (pp. 491-508). Hove, UK: Erlbaum.

Dempster, F. N. (1981). Memory span: Sources of individual and developmental differences. Psychological Bulletin, 89, 63-100.

Duncan, J., Burgess, P., & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. Neuropsychologia, 33, 261-268.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. Cognitive Psychology, 30, 257-303.

Engle, R. W., & Conway, A. R. A. (1998). Comprehension and working memory. In R. H. Logie & K. J. Gilhooly (Eds.), Working memory and thinking (pp. 67-91). Hillsdale, NJ: Erlbaum.

Engle, R. W., & Oransky, N. (1999). The evolution from short-term to working memory: Multi-store to dynamic models of temporary storage. In R. J. Sternberg (Ed.), The nature of cognition (pp. 515-555). Cambridge, MA: MIT Press.

Engle, R. W., Cantor, J., & Carullo, J. J. (1992). Individual differences in working memory and comprehension: A test of four hypotheses. Journal of Experimental Psychology: Learning, Memory, and Cognition, 18, 972-992.

Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. Journal of Educational Research, 84, 253 - 262.

Engle, R. W., Conway, A. R. A., Tuholski, S. W., & Shisler, R. J. (1995). A resource account of inhibition. Psychological Science, 6, 122-125.

Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), Models of working memory: Mechanisms of active maintenance and executive control (pp. 102-134). New York: Cambridge University Press.

Engle, R. W., Nations, J. K., & Cantor, J. (1990). Is "working memory capacity" just another name for word knowledge? Journal of Educational Psychology, 82, 799-804.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. Journal of Experimental Psychology: General, 128, 309-331.

Friedman, N. P., & Miyake, A. (2000). Differential roles for visuospatial and verbal working memory in situation model construction. Journal of Experimental Psychology: General, 129, 61-83.

Garner, W. R., Hake, H. W., & Eriksen, C. W. (1956). Operationism and the concept of perception. Psychological Review, 63, 149-159.

Goldinger, S.D., Kleider, H.M., Azuma, T., & Beike, D.R. (2003). "Blaming the victim" under memory load. Psychological Science, 14, 81-85.

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. Cognitive Psychology, 44, 339-384.

Hambrick, D.Z., Kane, M.J., & Engle, R.W. (in press). The role of working memory in higher-level cognition: Domain-specific versus domain-general perspectives. To appear in R. Sternberg & J. Pretz (Eds.), Cognition and Intelligence. New York: Cambridge University Press.

Handley, S. J., Capon, A., Copp, C., & Harper, C. (2002). Conditional reasoning and the Tower of Hanoi: The role of spatial and verbal working memory. British Journal of Psychology, 93, 501-518.

Hyde, T. S., & Jenkins, J. J. (1973). Recall of words as a function of semantic, graphic, and syntactic orienting tasks. Journal of Verbal Learning and Verbal Behavior, 12, 471-480.

Institute for Personality and Ability Testing (1973). Measuring Intelligence with Culture Fair Tests. Champaign, IL: Institute for Personality and Ability Testing.

Jensen, A. R. (1982). The chronometry of intelligence. In R. J. Sternberg (Ed.), Advances in the psychology of human intelligence (pp. 255-310). Hillsdale, NJ: Lawrence Erlbaum Associates.

Jensen, A. R. (1998). The g Factor: The Science of Mental Ability. Westport, CT: Praeger.

Jonides, J. & Smith, E. E. (1997). The architecture of working memory. In M. D. Rugg, (Ed). Cognitive neuroscience. Studies in cognition (pp. 243-276). Cambridge, MA: MIT Press.

Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 333-358.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. Psychonomic Bulletin and Review, 9, 637-671.

Kane, M. J., & Engle, R. W. (2003a). A role for working-memory capacity in task switching? Unpublished manuscript.

Kane, M. J., & Engle, R. W. (2003b). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. Journal of Experimental Psychology: General, 132, 47-70.

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. Journal of Experimental Psychology: General, 130, 169-183

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2003). The generality of working memory capacity: A latent-variable approach to verbal and visuo-spatial memory span and reasoning. Manuscript under review.

Kane, M. J., Poole, B., Tuholski, S. W., & Engle, R. W. (2003). Working-memory capacity and the executive control of visual search and enumeration. Unpublished manuscript.

Kiewra, K. A., & Benton, S. L. (1988). The relationship between information processing ability and notetaking. Contemporary Educational Psychology, 13, 33-44.

King, J. & Just, M. A. (1991). Indvidual differences in syntactic processing: The role of working memory. Journal of Memory and Language, 30, 580-602.

Klein, K., & Fiss, W. H. (1999). The reliability and stability of the Turner and Engle working memory task. Behavior Research Methods, Instruments and Computers, 31, 429-432.

Kosslyn, S. M., Cacciopo, J. T., Davidson, R. J., Hugdahl, K., Lovallo, W. R., Spiegel, W. R., & Rose, R. (2002). Bridging psychology and biology: The analysis of individuals in groups. American Psychologist, 57, 341-341.

Kyllonen, P. C. (1993). Aptitude testing inspired by information processing: A test of the four-sources model. Journal of General Psychology, 120, 375-405.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! Intelligence, 14, 389-433.

Kyllonen, P. C., & Stephens, D. L. (1990). Cognitive abilities as determinants of success in acquiring logic skill. Learning and Individual Differences, 2, 129-160.

Law, D. J., Morrin, K. A., & Pellegrino, J. W. (1995). Training effects and working memory contributions to skill acquisition in a complex coordination task. Learning and Individual Differences, 7, 207-234.

Logie, R. H. (1995). Visuo-spatial working memory. Hillsdale, NJ: Erlbaum.

Long, D. L., & Prat, C. S. (2002). Working memory and Stroop interference: An individual differences investigation. Memory & Cognition, 3, 294-301.

MacDonald, A. W., III, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000, June). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. Science, 288, 1835-1838.

MacLeod, C. M. (1998). Training on integrated versus separated Stroop tasks: The progression of interference and facilitation. Memory & Cognition, 26, 201-211.

MacLeod, C. M., & MacDonald, P. A. (2000). Interdimensional interference in the Stroop effect: Uncovering the cognitive and neural anatomy of attention. Trends in Cognitive Sciences, 4, 383-391.

MacLeod, C. M., Dodd, M. D., Sheard, E.D., Wilson, D. E. & Bibi, U. (in press). In opposition to inhibition. To appear in B. Ross (Ed.), The Psychology of Learning and Motivation. New York: Academic Press.

McNamara, D. S., & Scott, J. L. (2001). Working memory capacity and strategy use. Memory & Cognition, 29, 10-17.

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. Journal of Experimental Psychology: General, 130, 621-640.

Morrell, R. W., & Park, D. C. (1993). The effects of age, illustrations, and task variables on the performance of procedural assembly tasks. Psychology and Aging, 8, 389-399.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), Consciousness and self-regulation: Advances in research and theory (vol. 4, pp. 1-18). New York: Plenum Press.

O'Reilly, R. C., Braver, T. S., & Cohen, J. D. (1997, July). A biologically-based computational model of working memory. Paper presented at the "Models of Working Memory" Symposium, Boulder, CO.

Oberauer, K. (1993). Die Koordination kognitiver Operationen. Eine Studie zum Zusammenhang von 'working memory' und Intelligenz [The coordination of cognitive operations. A study of the relationship of working memory and intelligence]. Zeitschrift für Psychologie, 201, 57-84.

Oberauer, K., Süβ, H.-M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000). Working memory capacity – Facets of a cognitive ability construct. Personality and Individual Differences, 29, 1017-1045.

Oberauer, K., Süβ, H.-M., Wilhelm, O., & Wittmann, W. W. (2003). The multiple faces of working memory: Storage, processing, supervision, and coordination. Intelligence, 31, 167-193.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult lifespan. Psychology and Aging, 17, 299-320.

Perfetti, C. A., & Lesgold, A. M. (1977). Discourse comprehension and sources of individual differences. In M. A. Just & P. A. Carpenter (Eds.), Cognitive processes in comprehension (pp. 141-183). Hillsdale, NJ: Lawrence Erlbaum Associates.

Postman, L., Stark, K., & Fraser, J. (1968). Temporal changes in interference. Journal of Verbal Learning and Verbal Behavior, 7, 672-694.

Raven, J. C., Raven, J. E., & Court, J. H. (1998). Progressive Matrices. Oxford, England: Oxford Psychologists Press.

Richeson, J.A. & Shelton, J.N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. Psychological Science, 14, 287-291.

Roberts, R. J., Jr., Hager, L. D., & Heron, C. (1994). Prefrontal cognitive processes: Working memory and inhibition in the antisaccade task. Journal of Experimental Psychology: General, 123, 374-393.

Roediger, H. L. III, & Crowder, R. G. (1976). A serial position effect in recall of United States presidents. Bulletin of the Psychonomic Society, 8, 275-278.

Rogers, R. D., & Monsell, S. (1995). The cost of a predictable switch between simple cognitive tasks. Journal of Experimental Psychology: General, 124, 207-231.

Rosen, V. M., & Engle, R. W. (1997). The role of working memory capacity in retrieval. Journal of Experimental Psychology: General, 126, 211-227.

Rosen, V. M., & Engle, R. W. (1998). Working memory capacity and suppression. Journal of Memory and Language, 39, 418-436.

Rosen, V. M., Bergeson, J. L., Putnam, Harwell, A., & Sunderland, T. (2002). Working memory and apolipoprotein E: What's the connection? Neuropsychologia, 40, 2226-2233.

Salthouse, T. A. (1995). Differential age-related influences on memory for verbal-symbolic and visual-spatial information? Journal of Gerontology: Psychological Sciences, 50B, P193-P201.

Salthouse, T. A. (1996). The processing speed theory of adult age differences in cognition. Psychological Review, 103, 403-428.

Salthouse, T. A. (2001). A research strategy for investigating group differences in a cognitive construct: Application to ageing and executive processes. European Journal of Cognitive Psychology, 13, 29-46.

Shah, P., & Miyake, A. (1996). The separability of working memory resources for spatial thinking and language processing: An individual differences approach. Journal of Experimental Psychology: General, 125, 4-27.

Shute, V. J. (1991). Who is likely to acquire programming skills? Journal of Educational Computing Research, 7, 1-24.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. Journal of Experimental Psychology, 18, 643-662.

Süβ, H-M., Oberauer, K., Wittman, W. W., Wilhelm, O., & Schulze, R. (2002). Working-memory capacity explains reasoning ability – and a little bit more. Intelligence, 30, 261-288.

Swanson, H. L. (1996). Individual and age-related differences in children's working memory. Memory & Cognition, 24, 70-82.

Tuholski, S. W., Engle, R. W., & Baylis, G. C. (2001). Individual differences in working memory capacity and enumeration. Memory & Cognition, 29, 484-492..

Turley-Ames, K. J., & Whitfield, M. M. (2003). Strategy training and working memory task performance. Manuscript submitted for publication.

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? Journal of Memory and Language, 28, 127-154.

Underwood, B. J. (1975). Individual differences as a crucible in theory construction. American Psychologist, 30, 128-134.

West, R., & Alain, C. (2000b). Evidence for the transient nature of a neural system supporting goal-directed action. Cerebral Cortex, 10, 748-752.

Wilson, K. M., & Swanson, H. L. (2001). Are mathematics disabilities due to a domain-general or a domain-specific working memory deficit? Journal of Learning Disabilities, 34, 237-248.

Figure Captions

Figure 1: Measurement model of the working memory system (modified from Engle et al., 1999). The labels for James and Hebb refer to our observation that those two different perspectives led to the two different views of primary/STM as noted in Engle and Oransky (1999).

Figure 2: (A) Mean visual-search latencies on target-present trials for high and low WMC span subjects in regularly arranged, 4 x 4 search arrays. (B) Mean visual-search latencies on target-present trials for high and low span subjects for spatially irregular search arrays. For both panels, the less steep lines reflect latencies under relatively "automatic" search conditions and the upper two lines reflect latencies under relatively "controlled" search conditions. Display set size refers to the number of targets plus distractors in the arrays. Error bars depict standard errors of the means. HiAuto = high spans under automatic search conditions; HiCont = high spans under controlled search conditions; LoAuto = low spans under automatic search conditions; LoCont = low spans under controlled search condition; RT = response time; ms = milliseconds.

Figure 3: Path model for confirmatory factor analysis from Engle et al. (1999) showing the significant link between WMC and general fluid intelligence but the non-significant link between STM and gF.

Figure 4: Path model for confirmatory factor analysis from Engle et al. (1999) showing that, after variance common to the STM tasks and the WMC tasks was removed as Common, the correlation between the residual or left-over variance in WMC and gF was highly significant.

Figure 5: (A) Path model for confirmatory factor analysis consisting of a single WMC factor versus two domain-specific factors. Paths connecting manifest variables (boxes) to each other represent correlated error terms added to the model. (B) Path model for confirmatory factor analyses contrasting one- versus two-factor models, but with no correlated errors. In both panels, paths connecting latent variables (circles) to each other represent the correlations between the constructs, and numbers to the left of each manifest variable represent the loadings for each task onto the latent variable. WMC = working memory capacity; WMC-V = working memory capacity-verbal; WMC-S = working memory capacity-spatial.

Figure 6: Path model for structural equation analysis of the relation between working memory capacity and reasoning factors. Paths connecting manifest variables (boxes) to each other represent correlated error terms added to the model. Paths connecting latent variables (circles) to each other represent the correlations between the constructs. All paths are statistically significant. The numbers to the left of each WMC task represent the

loadings for each task onto the WMC factor. The numbers under the gF column on the right represent the factor loadings for each reasoning task onto the gF factor; the numbers under the Reas column represent the simultaneous factor loadings for each reasoning task onto either the verbal or spatial reasoning factors. WMC = working memory capacity; gF = general fluid intelligence; REA-V = reasoning-verbal; REA-S = reasoning spatial.

Figure 7: Path model for structural equation analysis of the relation between memory (short-term memory and working memory capacity) and reasoning factors. All paths are statistically significant, except the path (.16) from Storage-V to gF. The numbers under the Exec column on the left represent the factor loadings for each memory span task onto the ExecAttn factor; the numbers under the Stor column represent the simultaneous factor loadings for each memory span task onto either the verbal or spatial storage factor. The numbers under the gF column on the right represent the factor loadings for each reasoning task onto the gF factor; the numbers under the Reas column represent the simultaneous factor loadings for each reasoning task onto either the verbal or spatial reasoning factors. ExecAttn = executive attention; Storage-V = storage-verbal; Storage-S = storage-spatial; gF = general fluid intelligence; REA-V = reasoning-verbal; REA-S = reasoning spatial.

Figure 8: Percent error for high and low WMC subjects in prosaccade and antisaccade conditions. Error bars depict the standard errors of the means.

Figure 9: Mean error-rate interference effects for high and low WMC span participants in high congruency contexts (75% or 80%) across four experimental groups from Kane and Engle (2003). Interference effects were calculated by subtracting participants' mean baseline error rate from incongruent-trial error rate. Error bars depict standard errors of the means. E1 = Experiment 1; E2 = Experiment 2; E4a = Experiment 4a; E4b = Experiment 4b.

Figure 1

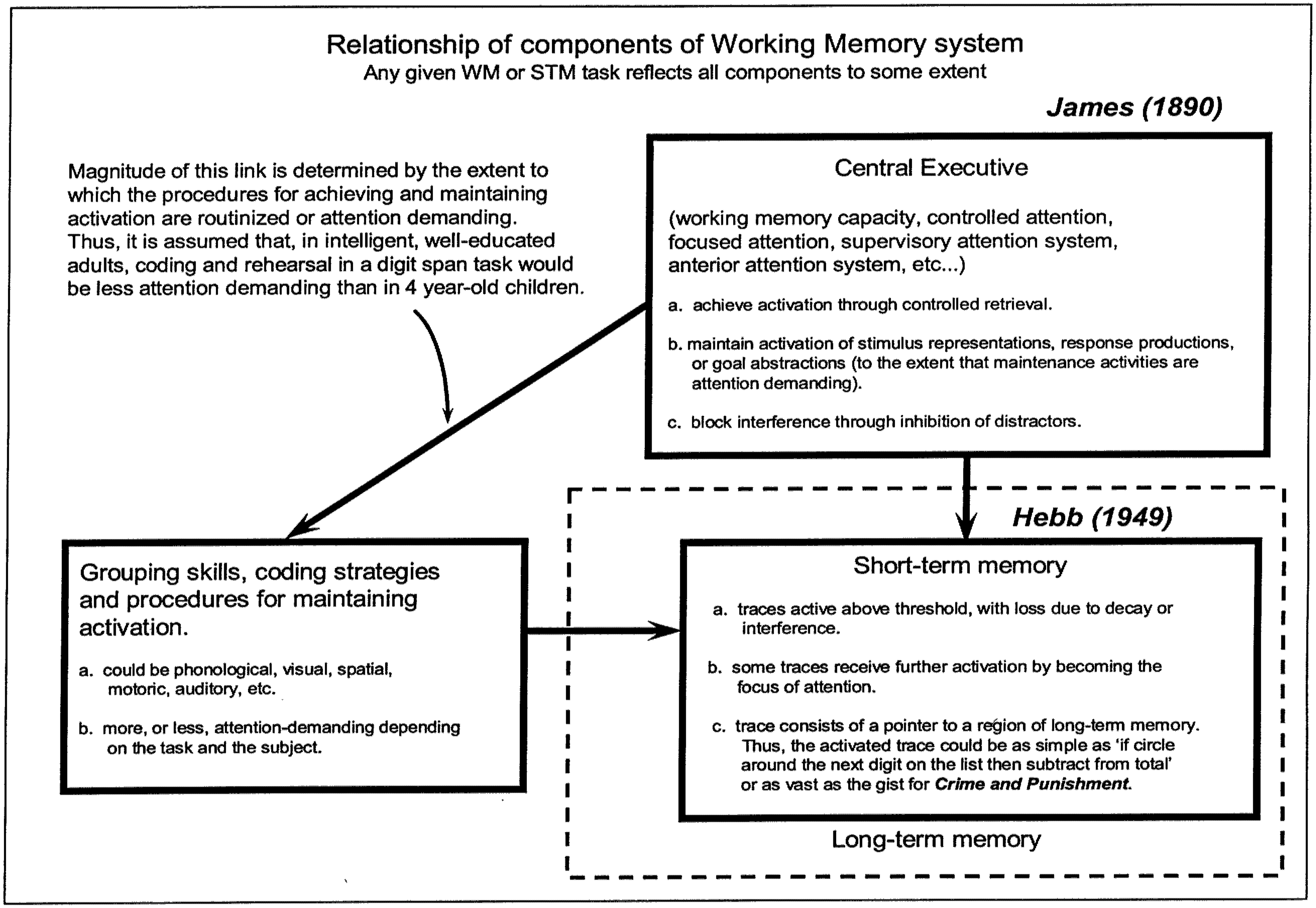

Relationship of components of Working Memory system
Any given WM or STM task reflects all components to some extent
James (1890)
Magnitude of this link is determined by the extent to which the procedures for achieving and maintaining activation are routinized or attention demanding. Thus, it is assumed that, in intelligent, well-educated adults, coding and rehearsal in a digit span task would be less attention demanding than in 4 year-old children.
Central Executive
(working memory capacity, controlled attention, focused attention, supervisory attention system, anterior attention system, etc...)
a. achieve activation through controlled retrieval.
b. maintain activation of stimulus representations, response productions, or goal abstractions (to the extent that maintenance activities are attention demanding).
c. block interference through inhibition of distractors.
Hebb (1949)
Grouping skills, coding strategies and procedures for maintaining activation.
a. could be phonological, visual, spatial, motoric, auditory, etc.
b. more, or less, attention-demanding depending on the task and the subject.
Short-term memory
a. traces active above threshold, with loss due to decay or interference.
b. some traces receive further activation by becoming the focus of attention.
c. trace consists of a pointer to a region of long-term memory. Thus, the activated trace could be as simple as 'if circle around the next digit on the list then subtract from total' or as vast as the gist for Crime and Punishment.
Long-term memory

Figure 2

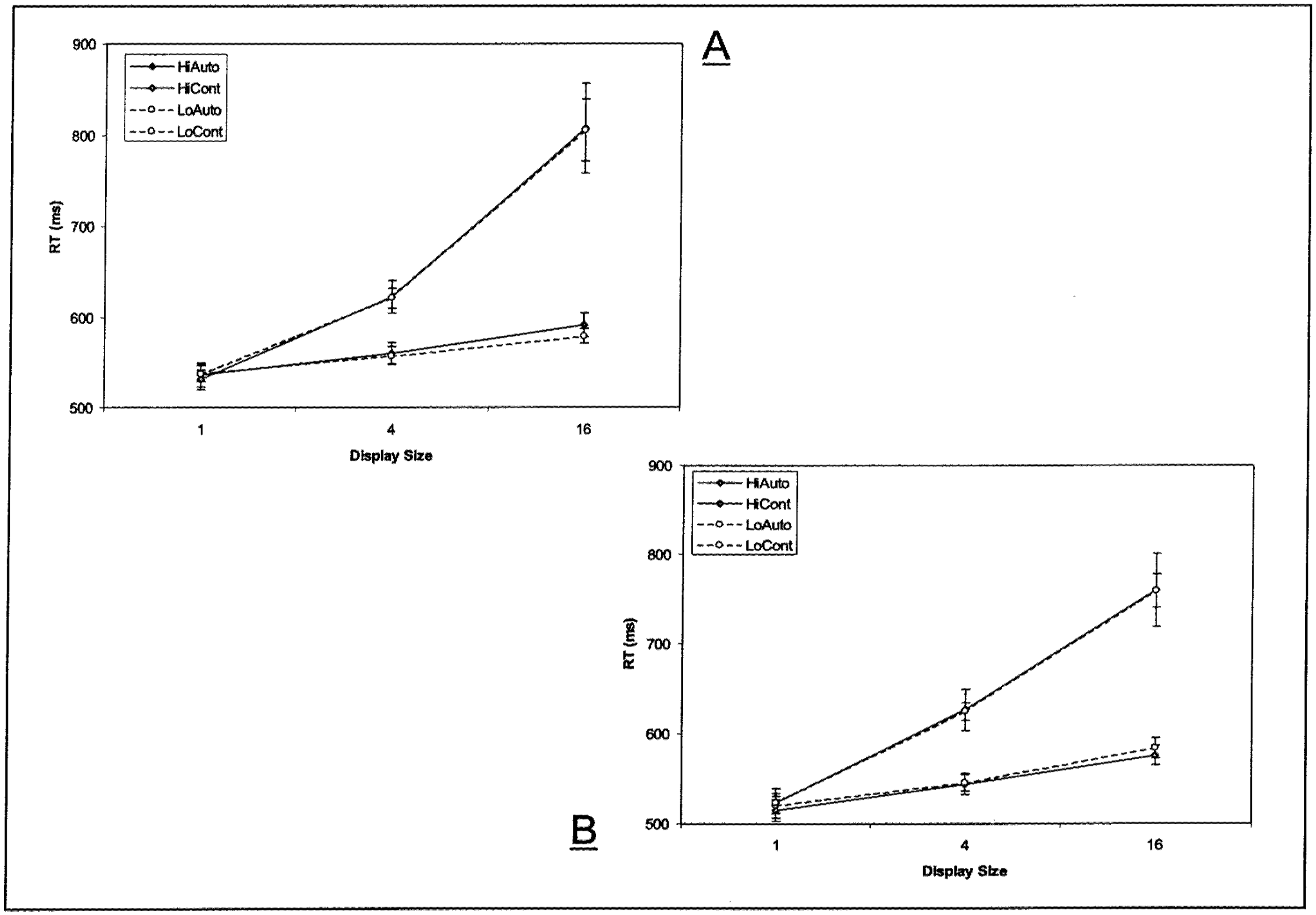
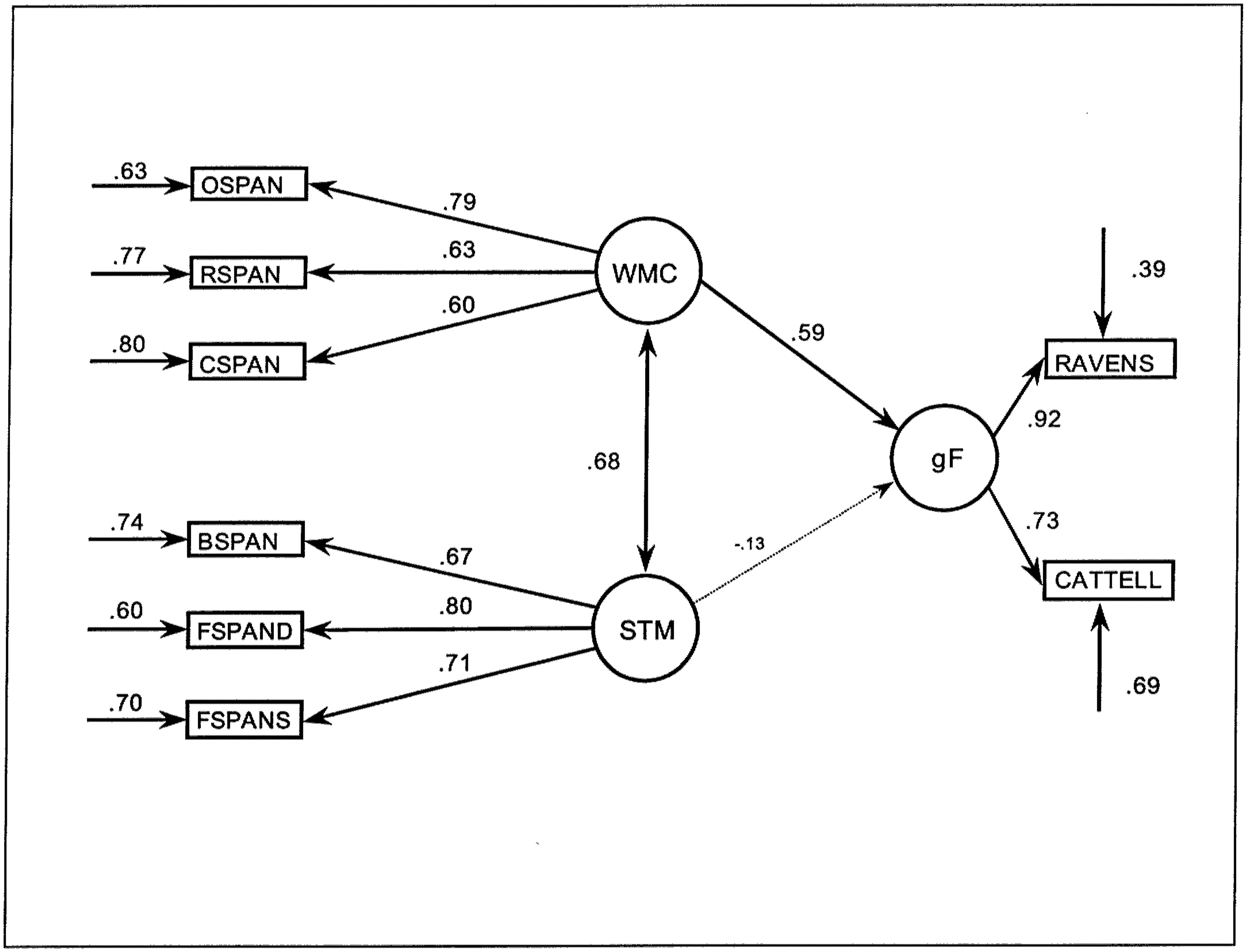
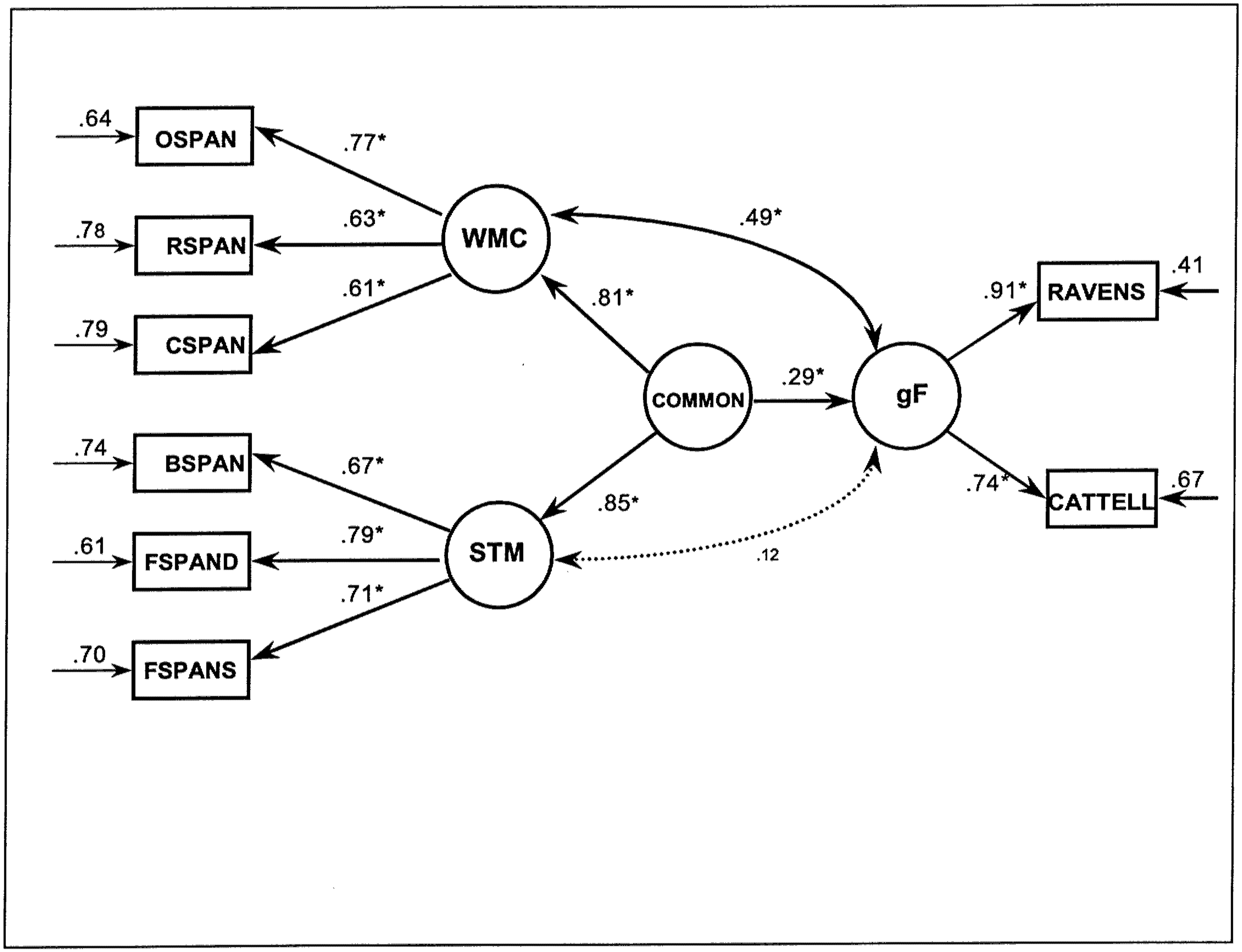

A
900
800
700
600
500
RT (ms)
HiAuto
HiCont
LoAuto
LoCont
1
4
16
Display Size
B
900
800
700
600
500
RT (ms)
HiAuto
HiCont
LoAuto
LoCont
1
4
16
Display Size

Figure 3



.63
OSPAN
.77
RSPAN
.80
CSPAN
.79
.63
.60
WMC
.68
.74
BSPAN
.60
FSPAND
.70
FSPANS
.67
.80
.71
STM
.59
-.13
gF
.92
.73
RAVENS
.39
CATTELL
.69

Figure 4



.64
OSPAN
.78
RSPAN
.79
CSPAN
.74
BSPAN
.61
FSPAND
.70
FSPANS
.77*
.63*
.61*
.67*
.79*
.71*
WMC
STM
.81*
.85*
COMMON
.29*
.49*
.12
gF
.91*
.74*
RAVENS
CATTELL
.41
.67

Figure 5

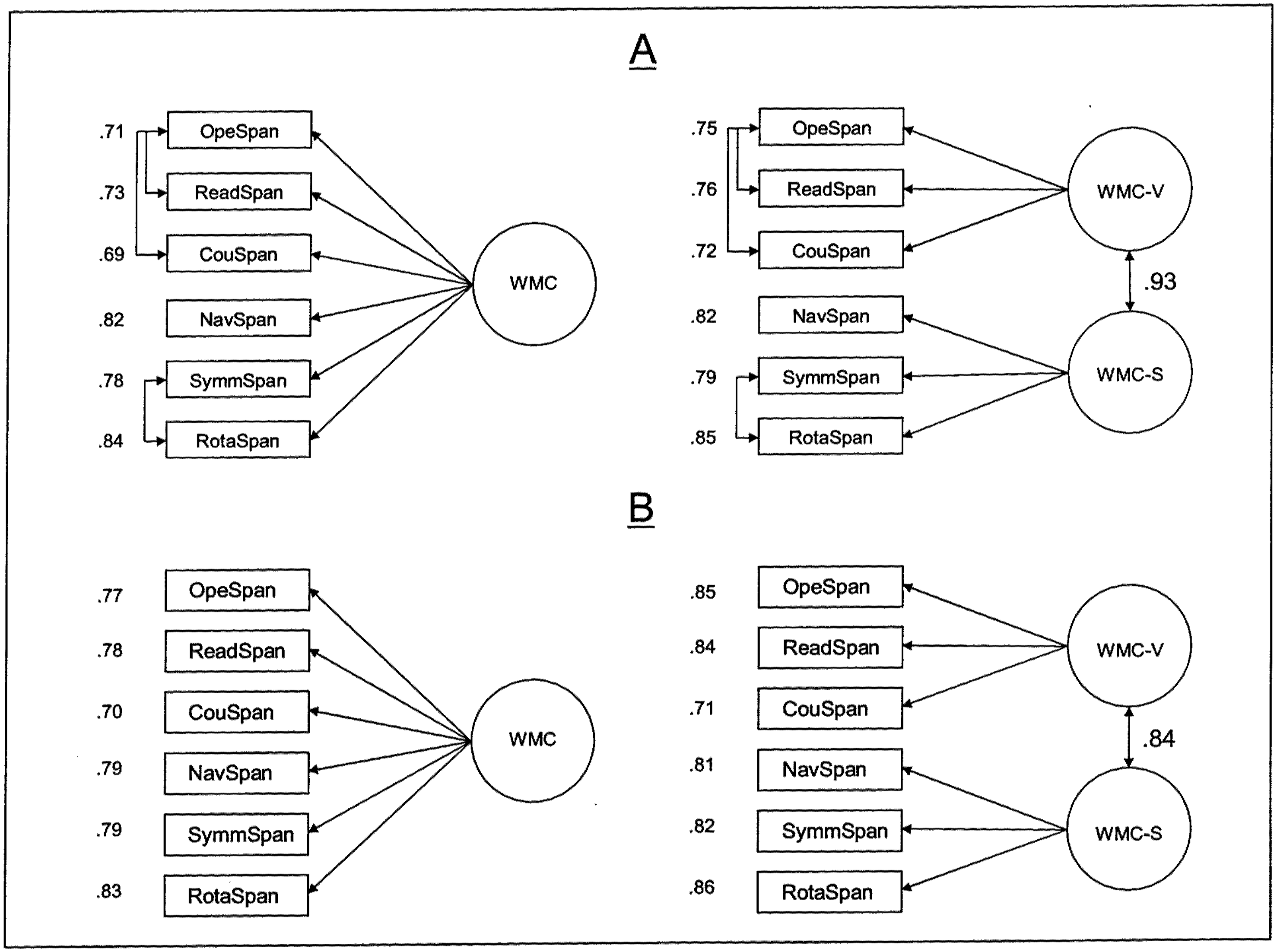
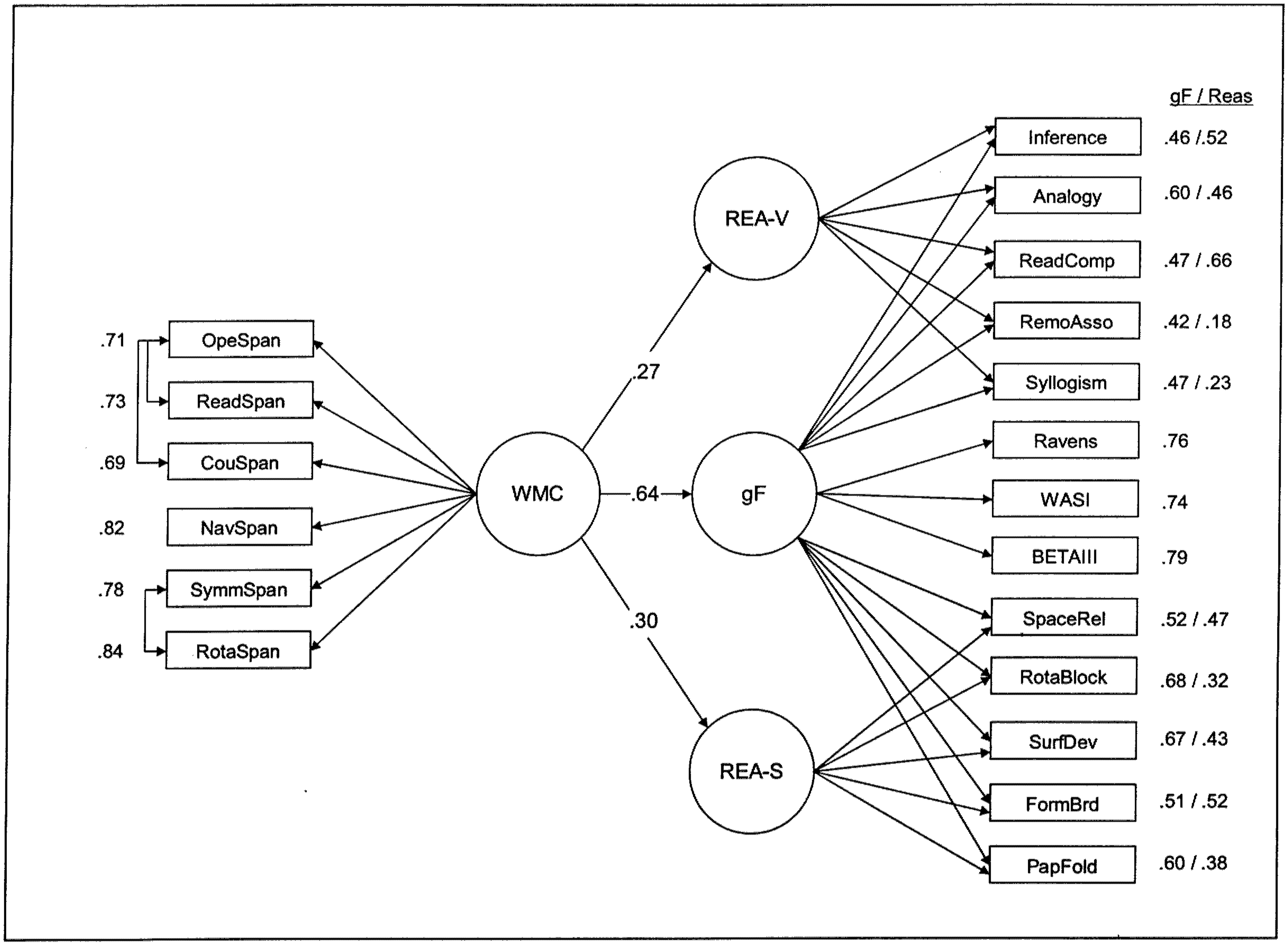
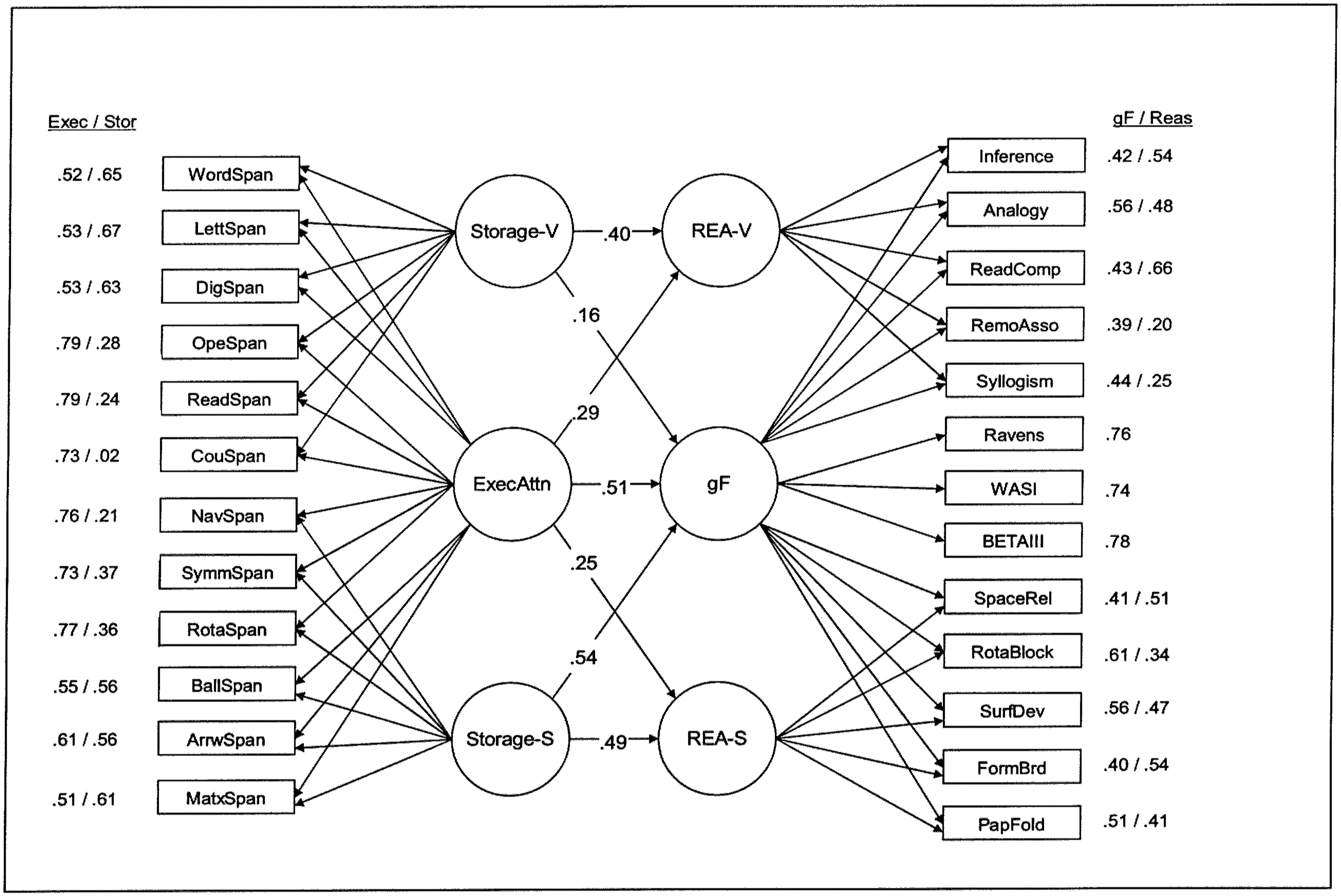

A
.71
OpeSpan
.73
ReadSpan
.69
CouSpan
.82
NavSpan
.78
SymmSpan
.84
RotaSpan
WMC
.75
OpeSpan
.76
ReadSpan
.72
CouSpan
.82
NavSpan
.79
SymmSpan
.85
RotaSpan
WMC-V
.93
WMC-S
B
.77
OpeSpan
.78
ReadSpan
.70
CouSpan
.79
NavSpan
.79
SymmSpan
.83
RotaSpan
WMC
.85
OpeSpan
.84
ReadSpan
.71
CouSpan
.81
NavSpan
.82
SymmSpan
.86
RotaSpan
WMC-V
.84
WMC-S

Figure 6



gF / Reas
Inference .46 /.52
Analogy .60 / .46
ReadComp .47 / .66
RemoAsso .42 / .18
Syllogism .47 / .23
Ravens .76
WASI .74
BETAIII .79
SpaceRel .52 / .47
RotaBlock .68 / .32
SurfDev .67 / .43
FormBrd .51 / .52
PapFold .60 / .38
REA-V
gF
REA-S
WMC
.27
.64
.30
.71 OpeSpan
.73 ReadSpan
.69 CouSpan
.82 NavSpan
.78 SymmSpan
.84 RotaSpan

Figure 7



Exec / Stor
.52 / .65
WordSpan
.53 / .67
LettSpan
.53 / .63
DigSpan
.79 / .28
OpeSpan
.79 / .24
ReadSpan
.73 / .02
CouSpan
.76 / .21
NavSpan
.73 / .37
SymmSpan
.77 / .36
RotaSpan
.55 / .56
BallSpan
.61 / .56
ArrwSpan
.51 / .61
MatxSpan
Storage-V
ExecAttn
Storage-S
.40
.16
.29
.51
.25
.54
.49
REA-V
gF
REA-S
gF / Reas
Inference
.42 / .54
Analogy
.56 / .48
ReadComp
.43 / .66
RemoAsso
.39 / .20
Syllogism
.44 / .25
Ravens
.76
WASI
.74
BETAIII
.78
SpaceRel
.41 / .51
RotaBlock
.61 / .34
SurfDev
.56 / .47
FormBrd
.40 / .54
PapFold
.51 / .41

Figure 8

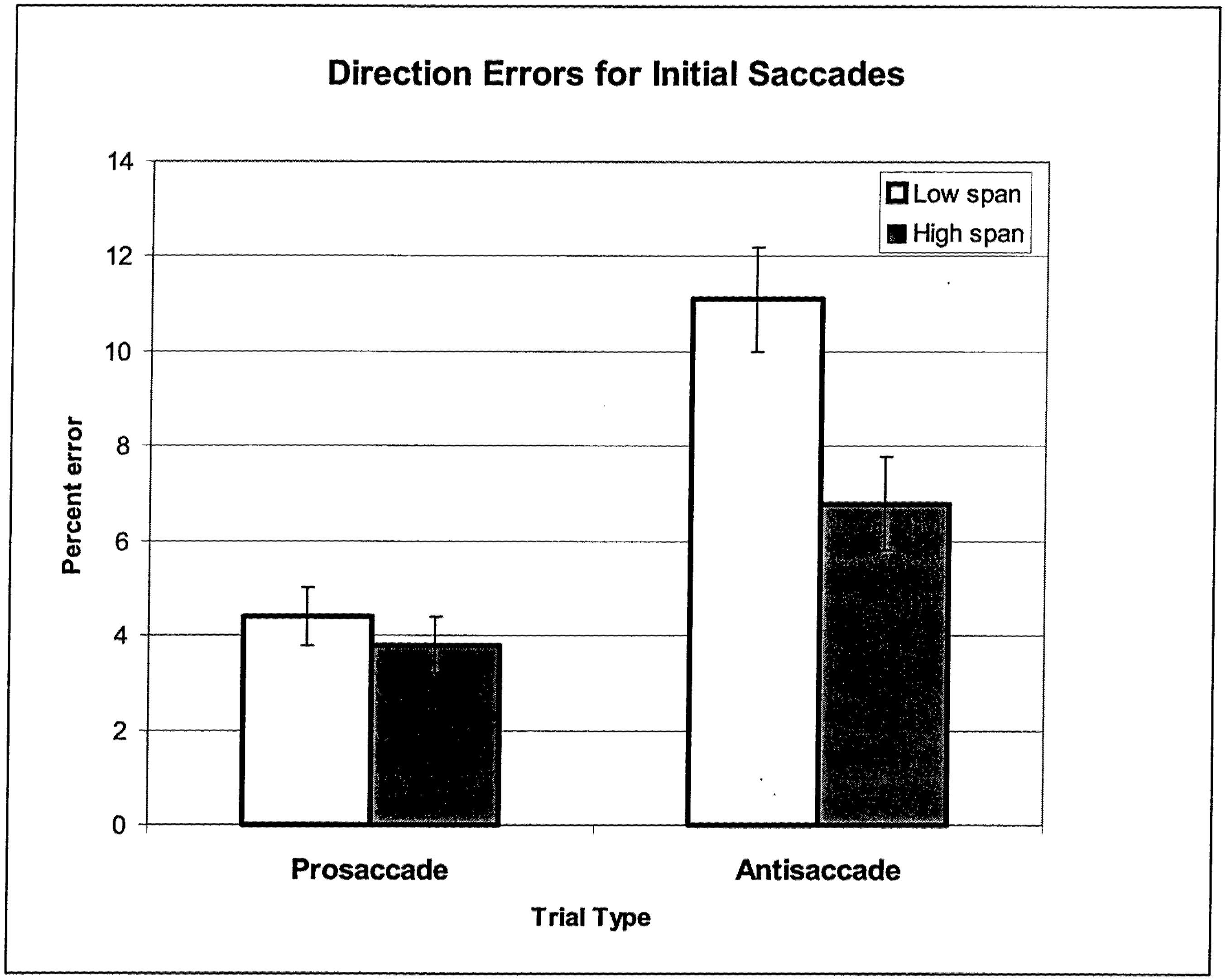
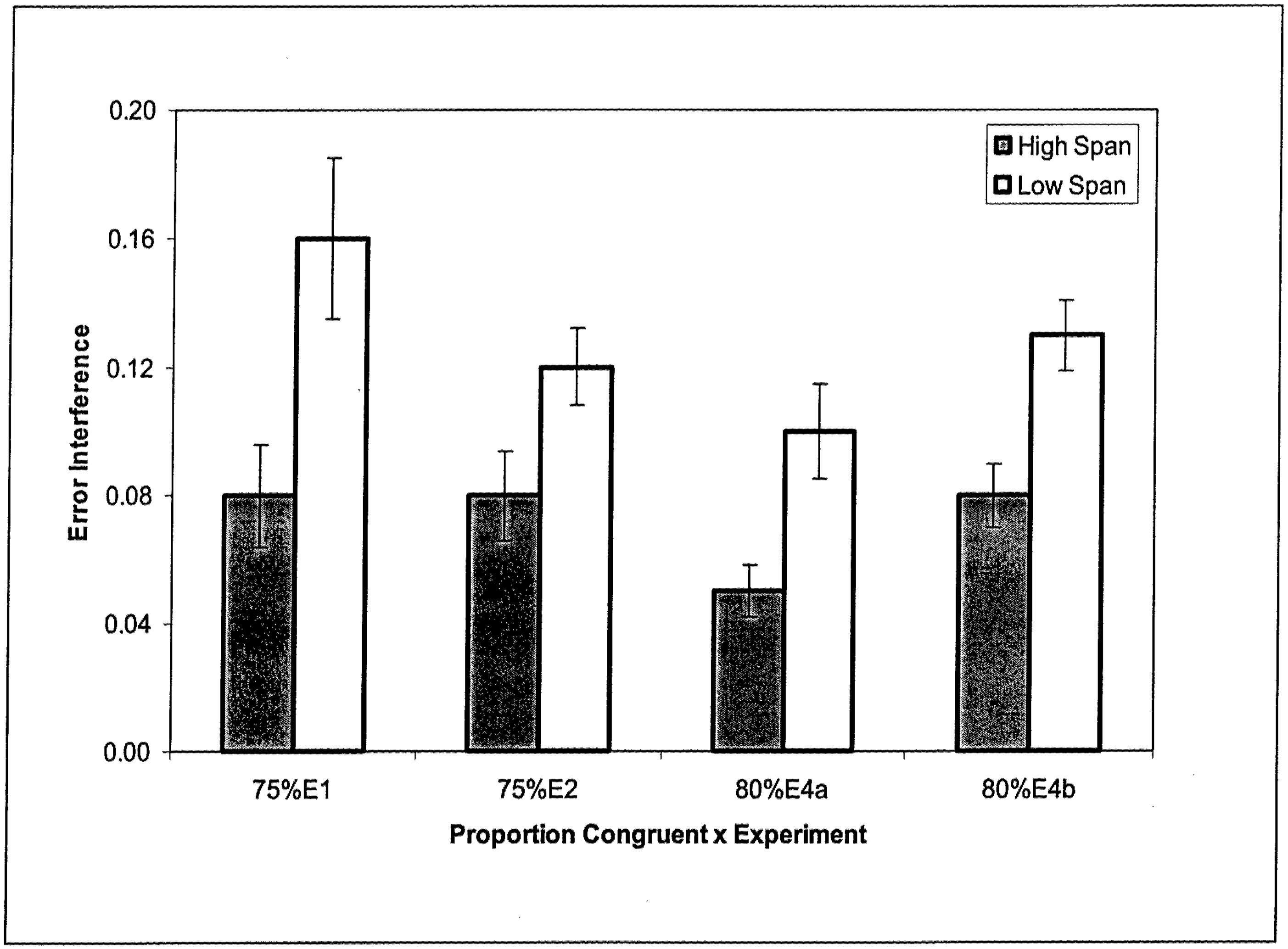

Direction Errors for Initial Saccades
Percent error
14
12
10
8
6
4
2
0
Low span
High span
Prosaccade
Antisaccade
Trial Type

Figure 9


0.20
0.16
0.12
0.08
0.04
0.00
Error Interference
High Span
Low Span
75%E1
75%E2
80%E4a
80%E4b
Proportion Congruent x Experiment

Scientific papers resulting from this grant

1. Engle, R.W. & Oransky, N. (1999). The evolution from short-term to working memory: Multi-store to dynamic models of temporary storage. In R. Sternberg (Ed.), The Nature of Cognition (pp. 514-555). Cambridge, MA: MIT Press.

2. Engle, R.W., Kane, M.J. & Tuholski, S.W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence and functions of the prefrontal cortex. In Miyake, A. & Shah, P. (Eds.), Models of working memory: Mechanisms of active maintenance and executive control ((pp.102-134). London:Cambridge Press.

3. Engle, R.W., Tuholski, S.W., Laughlin, J.E., & Conway, A.R.R. (1999). Working memory, short-term memory and general fluid intelligence: A latent variable approach. Journal of Experimental Psychology: General, 128, 309-331.

4. Kane, M.J., Conway, A.R.A., & Engle, R.W. (1999). What do working memory tests really measure? Behavioral and Brain Sciences, 22, 101-102.

5. Conway, A.R.A, Tuholski, S.W., Shisler, R.J. & Engle, R.W. (1999). The effect of memory load on negative priming: An individual differences investigation. Memory & Cognition, 27, 1042-1050.

6. Conway, A. R. A., Kane, M. J., & Engle, R. W. (1999). Is Spearman's g determined by speed or working memory capacity? Book review of Jensen on Intelligence-g-factor. Psycoloquy, 10(074).

7. Kane, M.J., & Engle, R.W. (2000). Working memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. Journal of Experimental Psychology: Learning, Memory, and Cognition, 26, 336-358.

8. Kane, M.J., Bleckley, M.K., Conway, A.R.A., & Engle, R.W. (2001). A controlled-attention view of working-memory capacity. Journal of Experimental Psychology: General, 130, 169-183.

9. Engle, R.W. (2001). What is working memory capacity? In Roediger, H. L., Nairne, J.S., Neath, I., Suprenant, A. M. The Nature of Remembering: Essays in Honor of Robert G. Crowder. Washington, D.C.: American Psychological Association Press.

10. Tuholski, S. W., Engle, R. W., Baylis, G. C. (2001). Individual differences in working memory capacity and enumeration. Memory & Cognition, 29, 484-492.

11. Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2001). The role of working memory in learning disabilities. Issues in Education, 7, 87-95.

12. Engle, R. W. (2002). Working memory capacity as executive attention. Current Directions in Psychological Science, 11, 19-23.

13. Hambrick, D. Z. & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity and age on cognitive performance. Cognitive Psychology, 44,339-387.

14. Kane, M.J. & Engle, R.W. (2003). The Role of Prefrontal Cortex in Working-Memory Capacity, Executive Attention, and General Fluid Intelligence. Psychonomics Bulletin and Review, 9, 637-671..

15. Kane, M.J. & Engle, R.W. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. Journal of Experimental Psychology: General, 132, 47-70..

16. Hambrick, D. Z. & Engle, R. W. (2003). The role of working memory in higher level cognition. In Davidson, J. & Sternberg, R. The nature of problem solving. MIT Press.

17. Bleckley, M. K., Durso, F. T., Crutchfield, J. M., Engle, R., & Khanna, M. (in press). Individual differences in working memory capacity predict visual attention allocation. Psychonomic Bulletin and Review.

18. Hambrick, D. Z., Kane, M. J., & Engle, R. W. (in press) The Role of Working Memory in Higher-Level Cognition: Domain-Specific vs. Domain-General Perspectives. In R. J. Sternberg & J. Pretz, Eds. Cognition and Intelligence, Cambridge University Press.